Psychometrika

COPYRIGHT, PSYCHOMETRIC CORPORATION, 1937

CONTENTS

STUDIES IN THE LEARNING FUNCTION - - - - - 1
L. E. WILEY AND A. M. WILEY

MATRIX APPROXIMATIONS AND SUBSPACE FITTING - 21
GALE YOUNG

THE ESSENTIAL CONDITIONS OF MEASUREMENT - 27
DOUGLAS E. SCATES

THE RELATION BETWEEN RESULTS OBTAINABLE WITH RAW AND CORRECTED CORRELATION COEFFICIENTS IN MULTIPLE FACTOR ANALYSIS - - - - - - - - - - - - - 35
MERRILL ROFF

THE BI-FACTOR METHOD - - - - - - - - - - - 41
KARL J. HOLZINGER AND FRANCES SWINEFORD

A SEVEN-DECIMAL TABLE OF THE AREA (α) UNDER THE UNIT NORMAL CURVE, FOR ABSCISSAE EXPRESSED IN TERMS OF *P.E.* - - 55
RUTH H. KRAUSE AND HERBERT S. CONRAD

ABSTRACTS OF PAPERS TO BE PRESENTED AT THE DISTRICT MEETING OF THE PSYCHOMETRIC SOCIETY, THE UNIVERSITY OF CHICAGO, ON SATURDAY, APRIL 3, 1937 - - - 67

VOLUME TWO MARCH 1937 NUMBER ONE

PSYCHOMETRIKA—VOL. 2, NO. 1
MARCH, 1937

STUDIES IN THE LEARNING FUNCTION*

L. E. WILEY AND A. M. WILEY

Ohio Wesleyan University

The error function of the theoretical learning curve developed by Thurstone has been fitted to maze data from rats that have been subjected to cerebral cortical insult. There is a high degree of relationship between the number of errors necessary to complete learning as predicted from the curve and the experimental measure of learning, total-errors-minus-errors-first-trial. The procedure provides a method of analysis for individual learning records and it may be used to shorten the training period. The goodness of fit indicates that the underlying logic of the learning curve is plausible.

1. An Empirical Test of Thurstone's Theoretical Learning Curve

The purpose of the series of studies of which this paper is a part is to test empirically the theoretical learning curve developed by L. L. Thurstone ('30, '33).

The data were collected in collaboration with K. S. Lashley. A full description of the animals and procedure appears in Lashley and Wiley, "Mass Action in Relation to the Number of Elements in the Problem to be Learned." (1933)

Five mazes of the Lashley-type were used, Mazes I, II, III, and IV differed only in the number of culs-de-sac. Maze V was like Maze II except that the direction of running was reversed. One hundred and twenty-seven rats with varying amounts of brain lesion and sixty normal animals were trained. The animals were divided into four groups, approximately equal as to the percentage and locus of cerebral lesion. Each group was trained on two mazes, I, II, III, or IV and on Maze V. Thus, Group I was trained on Maze I and on Maze V; Group II on II and V, and so on. All of the animals were trained on Maze V. They were trained on the variable maze first.

For the present analysis we chose the 30 animals, with brain lesions, that were trained on Maze I and Maze V. Maze I is a 4 cul-de sac maze and V is an 8 cul-de-sac maze. The animals were, whenever possible, trained to complete learning as defined by the fact that they ran ten consecutive errorless runs. If, however, they ran in

*A grant-in-aid from the National Research Council has made the present analysis possible. We are indebted to Professor L. L. Thurstone for aid and encouragement.

Maze I to 150 trials without reaching this point, training was stopped. In Maze V they were trained to 100 trials in the case that they did not reach the prescribed degree of learning. So that, in a number of cases, the animals were trained for fifty days, five trials a day. The labor involved in such a procedure is tremendous.

If we can find some sort of theoretical curve which fits our data in the first part and which, when projected, gives us results as good as obtained by a long training period, we can decrease this labor markedly. With this in mind we have fitted Thurstone's theoretical curve to these data.

In the former study our data were analyzed using as a measure of learning the criterion, total-errors-minus-the-errors-on-the-first trial. We chose this value because we found that it had a higher inter-maze correlation than did trials, time, or total errors. Such a procedure is obviously not the best in that some of the animals never did reach ten consecutive errorless runs. By fitting a theoretical curve to the learning record we can determine when the animal would have mastered the problem if given sufficient time.

Our data are especially well adapted to a test of the learning curve because the wide range of lesion produced a wide range of ability in the animals.

The theoretical learning curve is of the form

$$u = \frac{\sqrt{m}}{aK} - \frac{\sqrt{m}}{K} \cdot \frac{u}{R} , \tag{1}$$

where u represents the accumulated errors, R is the number of trials, K the learning constant of the animal, m the difficulty of the maze, and a is an arbitrary constant that can be absorbed in m or in the unit of measurement. In our analysis the curve appears in the form

$$u = A + \frac{BR}{C + R} , \tag{2}$$

where A, B, and C are constants. The difference lies in the fact that Thurstone's curve represents the situation in which the origin of the system of coordinates lies at the place where *learning begins*. In our curve the origin lies at the point where *training* begins. We do not actually know when the rat begins to learn. Equation (2) may be derived from (1) as follows:

Writing (1) explicitly for u we have

$$u = \frac{\sqrt{m}\,R}{a(KR + \sqrt{m})} . \tag{3}$$

But in the experimental situation the origin lies at the point where training begins. Let R' represent the number of trials measured from this point and u' the number of errors.
Then

$$u' = u + e \ , \quad R' = R + r \ . \tag{4}$$

Here e represents the number of errors the rat makes before he begins to learn and r represents the number of trials he has to run before he begins to learn. The values of e and r represent the adjustment period of the rat. They may be either positive or negative, depending on the rapidity of adjustment not only to the learning situation but to the period of handling which takes place before training begins.

Substituting (4) in (3) we have

$$u' - e = \frac{\sqrt{m}(R' - r)}{a[K(R' - r) + \sqrt{m}]} \ . \tag{5}$$

Dividing the numerator and denominator of the righ hand member by aK

$$u' - e = \frac{\frac{\sqrt{m}}{aK}(R' - r)}{R' - r + \frac{\sqrt{m}}{K}} \ .$$

Let

$$C \equiv \frac{\sqrt{m}}{K} - r \ . \tag{6}$$

Then

$$u' - e = \frac{\frac{\sqrt{m}}{aK}(R' - r)}{C + R'} \ .$$

Transposing and dividing the fraction by $(C + R')$,

$$u' = e + \frac{\sqrt{m}}{aK}\left[-\frac{r}{C} + \frac{(1 + \frac{r}{C})R'}{C + R'}\right].$$

Then let

$$A \equiv c - \frac{\sqrt{m}}{aK} \cdot \frac{r}{C} \text{ and } B \equiv \frac{\sqrt{m}}{aK}(1 + \frac{r}{C}) \tag{7}$$

and

$$u' = A + \frac{B\,R'}{C + R'} \cdot \tag{8}$$

Dropping the primes the equation takes the form of (2) in which A, B, and C are constants, u represents the accumulated errors and R the number of trials, measured from the origin of training.

The attribute of the theoretical curve which is most like the former criterion of learning, total-errors-minus-errors-on-the-first-trial, is the horizontal asymptote when u is represented by the axis of ordinates and R by the abscissa.

Examining (8) for the value of the asymptotes:

$$\begin{aligned} u' &= A + B && \text{as } R' \text{ approaches infinity,} \\ R' &= -C && \text{as } u' \text{ approaches infinity.} \end{aligned} \tag{9}$$

Likewise examining (1) for its limits:

$$\begin{aligned} u &= \frac{\sqrt{m}}{aK} && \text{as } R \text{ approaches infinity,} \\ R &= -\frac{\sqrt{m}}{K} && \text{as } u \text{ approaches infinity.} \end{aligned} \tag{10}$$

But the relationship expressed in (4) holds in the limits. Therefore

$$A + B = \frac{\sqrt{m}}{aK} + e \ , \tag{11}$$

and

$$-C = -\frac{\sqrt{m}}{K} + r \ , \tag{12}$$

which is the substitution already make in (6).

Equation (11) represents the substitution (7) by the fact that

$$A + B = e - \frac{\sqrt{m}}{aK} \cdot \frac{r}{C} + \frac{\sqrt{m}}{aK}\left(1 + \frac{r}{C}\right) = e + \frac{\sqrt{m}}{aK} \ ,$$

which is an identity.

Plates 1 to 30 show how well the theoretical curve fits the experimental data. The number of trials, R', is plotted on the horizontal axis, the accumulated errors, u', on the vertical axis. The continuous curve is the result of fitting the theoretical equation to the experimental data.

In fitting our data we used the method of selected points. Referring to (1) we readily see that the equation is linear in u and $\frac{u}{R}$. But since the experimental data rarely go through the origin of learning we may change our coordinates so that

$$u = u_1 - u' \ , \quad R = R_1 - R' \ ,$$

where R' and u' are measured from the origin of training and (R_1, u_1) is some point which appears from the data to fall on the experimental curve. (1) then takes the form

$$u_1 - u' = \frac{\sqrt{m}}{aK} - \frac{\sqrt{m}}{K} \cdot \frac{u_1 - u'}{R_1 - R_1} \ ,$$

or

$$u_1 - u' = b - a \frac{u_1 - u'}{R_1 - R'} \ , \qquad (13)$$

when a and b are constants representing the slope and intercept of the straight line.

We may, then, choose any other two points which apparently fall on the continuous curve and solve for the values of a and b by writing two simultaneous linear equations of form (13). A typical solution is shown in Table I, which represents the case for animal No. 13 on Maze I.

TABLE I

Method of Fitting Learning Curve to Experimental Data
No. 13, Maze I

u	R	142-u	66-R	$\frac{142-u}{66-R}$
17	3	125	63	1.984
90	26	52	40	1.300
142	66			

(13)

$125 = b - a(1.984)$

$52 = b - a(1.300)$

Subtracting

$73 = -a(0.684)$

$a = -106.73$

Solving (13) for b

$b = 52 - (106.73)(1.300)$

$b = -86.75$

(14)

$$u = \frac{-86.75(66-R)}{106.73 - (66-R)} + 142$$

$$u = \frac{-5725.50 + 86.75R}{40.73 + R} + 142$$

$$u = -140.57 + \frac{227.32R}{40.73 + R} + 142$$

$$u = 1.430 + \frac{227.32R}{40.73 + R}$$

R	u(calc)
3	17
10	46
20	76
30	98
50	127
66	142

Letting $R_1 = 66$, $u_1 = 142$, and choosing the other two points $R' = 3$, $u' = 17$, and $R' = 26$, $u' = 90$, we write our two simultaneous equations. Solving for a and b we find that $a = -106.73$, $b = -86.75$.

Writing (13) explicitly for u'

$$u' = \frac{-b(R_1 - R')}{a + (R_1 - R')} + u_1 .$$

Choosing values of R' that are fairly well spaced on the curve we solve for the corresponding theoretical values of u' by substitution in the theoretical curve. We then draw the curve through these values. In terms of (8) we have $A = 1.430$, $B = 227.32$, $C = 40.73$.

Table II represents the calculated values of A, B, and C for all

TABLE II

Calculated Constants of the Theoretical Learning Curve

Animal No.	Maze I			Maze V		
	A	B	C	A	B	C
1	-6.06	47.25	7.46	86.57	-23.78	-1.05
2	-8.70	53.34	3.43	9.41	32.32	18.23
3	11.31	43.47	26.56	38.41	116.44	5.13
4	55.51	-15.82	-2.08	-92.59	235.14	2.33
5	-8.55	65.95	8.32	6.00	86.27	5.09
6	137.48	-60.90	-3.21	42.97	192.93	5.43
7	-2.72	24.47	2.65	43.09	56.57	10.65
8	11.26	46.26	4.27	227.39	-184.17	-0.56
9	-47.00	106.60	3.60	6.09	153.94	3.61
10	-27.71	77.45	5.47	33.06	155.44	38.67
11	-19.78	54.72	2.25	-119.36	520.80	3.60
12	-223.28	668.63	5.60	230.80	508.92	17.88
13	1.43	227.32	40.73	14.20	338.55	43.17
14	-20.62	158.24	8.06	74.24	132.94	32.36
15	304.03	-190.99	-1.16	32.98	90.49	36.18
16	-3.46	42.91	2.91	119.38	-24.26	-10.88
17	-15.67	83.28	5.33	42.38	38.16	6.87
18	-122.52	370.14	13.80	34.29	750.45	24.95
19	12.42	85.09	10.69	-0.68	141.58	33.33
20	9.40	245.54	23.93	105.03	334.58	9.36
21	8.56	354.40	28.59	-26.33	636.56	63.17
22	-19.20	67.00	3.78	-11.83	331.66	12.73
23	76.02	548.37	39.89	408.00	559.48	18.37
24	-51.95	362.52	10.06	-231.75	778.79	2.71
25	22.91	181.69	37.66	76.77	228.94	29.66
26	34.99	1934.68	289.45	131.90	601.39	41.98
27	-871.45	1547.31	7.05	47.30	1106.92	15.67
28	-55.40	337.86	4.19	-37.31	413.74	18.58
29	2523.76	-1416.80	-10.54	82.12	9936.88	602.81
30	-177.14	1814.55	27.14	425.88	1063.14	15.43

of the animals for each maze. Thus the equation for animal 24 on Maze V is

$$u' = -231.75 + \frac{778.79\,R'}{2.71 + R'}.$$

Table III shows the horizontal asymptotes of the theoretical curve on Maze I and on Maze V. It also shows the percentage of lesion, and the former criterion of learning, total-errors-minus-er-

TABLE III

Comparison of Horizontal *Asymptote* of the Theoretical Curve with the criterion *Errors-minus-errors-on-first-trial*

Animal No.	Percent lesion	Horizontal Asymptotes		Errors-minus-errors-1st-trial	
		Maze I	Maze V	Maze I	Maze V
1	3.0	41.2	62.8	18	27
2	3.0	44.6	41.7	35	25
3	3.3	54.8	154.9	23	106
4	4.9	39.7	142.6	35	97
5	5.7	57.4	92.3	46	50
6	9.8	76.6	235.9	60	165
7	9.9	21.8	99.7	13	61
8	10.9	57.5	43.2	31	33
9	13.4	59.6	160.0	37	129
10	14.1	49.7	188.5	40	71
11	14.3	34.9	401.4	44	275
12	15.2	445.4	739.7	561	549
13	17.9	228.2	352.8	130	253
14	18.4	137.6	207.2	100	108
15	19.1	113.0	123.5	88	64
16	19.3	39.5	95.1	33	72
17	19.6	67.6	80.5	57	30
18	20.7	247.6	784.7	280	648
19	21.0	97.5	140.9	61	70
20	21.9	254.9	439.6	249	324
21	24.6	363.0	610.2	367	353
22	25.2	47.8	319.8	38	272
23	28.3	624.4	967.5	566	831
24	31.4	310.6	547.0	258	478
25	31.9	204.6	305.7	147	220
26	34.7	(1969.7)	733.3	563	451
27	36.6	675.9	1154.2	609	1044
28	36.8	282.5	376.4	226	300
29	41.7	1107.0	(10019.0)	1330	1448
30	65.3	1637.4	1489.0	1356	1462

rors-first-trial, for both mazes. The theoretical asymptotes were calculated from the value of u' given in (9). Thus for No. 24 on Maze V the asymptote will be (—231.75 + 778.79) or 547.0.

Cases 26 and 29 have not been included in the comparisons between mazes. Upon inspection of the experimental data in the plates it can readily be seen that the learning curves for 26 on Maze I and 29 on Maze V are practically straight lines. In other words the rats did not learn in these cases. It is, however, possible to get the constants of the learning curve from these date because of the fact that the straight line is a limiting case of our curve.

Figures I and II show the relationship between the asymptotes of the theoretical curves and the criterion, total-errors-minus-errors-first-trial. Figure III shows the relationship between the calculated asymptotes on Maze I and Maze V.

Plates 1 to 30 are evidence that in the majority of cases the theoretical curve fits the experimental data.

We have a few marked exceptions. In No. 6, Maze I we find a situation in which the curve is discontinuous. The first part of the period of training is linear. The rat does not begin to eliminate errors until the eighth trial. The first few points are not represented in the learning situation at all. They belong to the initial adjustment period.

Another type of deviation is represented by No. 12, Maze I. Here again the curve becomes discontinuous. Apparently the learning process changes. These changes may be due to the physical condition of the animal or to a difference in motivation.

We have still another type of difference shown in No. 18, Maze I. In this case the rat learns efficiently for a long period. Then he suddenly ceases to eliminate errors, making practically the same number of errors in each trial, and his learning curve becomes a straight line. Upon inspection of the actual training records we find that these animals are making the same errors over and over again. They have learned something, but they have failed to learn the problem as it is set to them by the experimenter.

In all cases in which such differences appear, we have fitted that part of the curve in which learning takes place and not the other parts.

Figures I and II show that there is a high degree of relationship between the calculated asymptotes and the former criterion of total-errors-minus-errors-first-trial. In spite of the fact that we have ignored such variations as we have discussed we find results of the same

magnitude from the calculated asymptotes that we get from the former criterion. We have, therefore, established the fact that we can get results from this method that are as good as those we get from long training.

It is possible to use the theoretical learning curve to great advantage in training animals on a maze problem in which the learning proves to be very slow. The animals can be trained until they show a smooth and continuous curve. The data can then be fitted with the theoretical learning curve, and the point at which learning is completed can be predicted. We suggest that in maze studies in which the method is used the experimenter plot his cumulated error curve as he does his training. In this manner he can determine when the curve becomes smooth enough to justify a cessation of training. In most cases there will be a great saving of time and the results will be just as reliable as they would be if he trained to some arbitrary criterion of learning, as shown by the linear relationship between the former criterion and the calculated asymptotes.

Figure III shows that the relationship between the asymptotes on the two mazes is linear. The two points which deviate most from the line are cases in which we have "stereotyped" behavior. This linearity shows that what the rat does on one maze he tends to do on another problem of the same sort. In other words, the learning constant of the animal remains the same from one problem to the next.*

The fact that this particular relationship is linear is very significant. It provides evidence that the logic underlying the theoretical development of the learning curve is reasonable. It establishes the internal consistency of the problem.

Conclusions:

1. The theoretical learning curve developed by Thurstone fits learning data collected from the Lashley-type maze. The horizontal asymptote of the curve has a high degree of relationship with the criterion, total-errors-minus-errors-on-the-first-trial.

2. It is possible to use this curve to predict the point at which learning is complete. It is not necessary to train animals to complete learning, if they are trained until a continuous function appears.

3. The ability of the animal remains constant from one maze to the next which shows that the underlying logic of the learning curve is plausible.

*A paper is in preparation in which we have a solution for the learning constants of the animals and for the relative difficulties of the two mazes.

Figure I. Relationship between the calculated horizontal asymptote and total - errors - minus - errors - on - the-first-trial for Maze I.

Figure II. Relationship between the calculated horizontal asymptote and total - errors - minus - errors - on - the-first-trial for Maze V.

Figure III. Relationship between the calculated horizontal asymptotes for Maze I and for Maze V.

REFERENCES

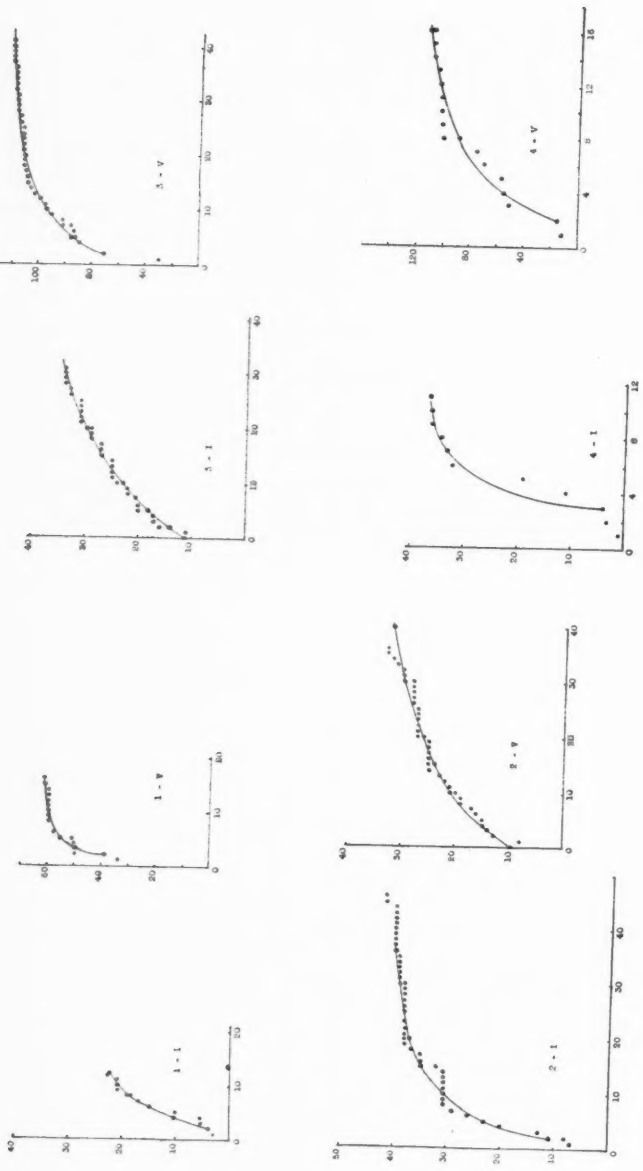
LASHLEY, K. S. AND L. E. WILEY, "Studies of cerebral function in learning. IX. Mass action in relation to the number of elements in the problem to be learned;" J. Comp. Neur., v. 57, 1933, no. 1, Feb. 15.

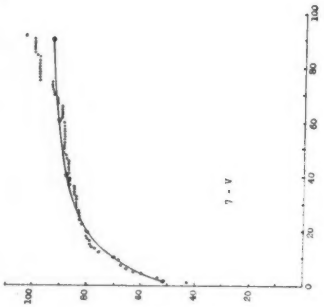
THURSTONE, L. L., 1930. "The learning function," Journal of General Psychology, 1930, 3, 469-493.

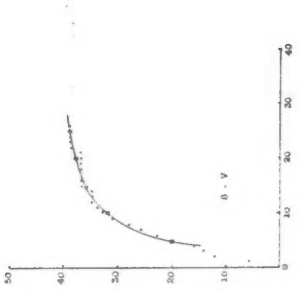
THURSTONE, L. L., 1933. "The error function in maze learning." Journal of General Psychology, 1933, 9, 288-301.

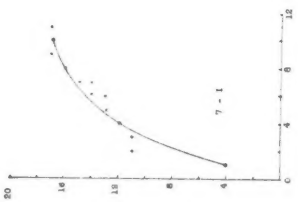
PLATES 1 to 30

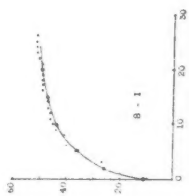
Plates 1 to 30 show the cumulative error curves for each animal on both mazes. The number of the animal is in Arabic notation, the number of the maze in Roman numerals. The horizontal axis represents the number of trials and the vertical axis the cumulative errors.

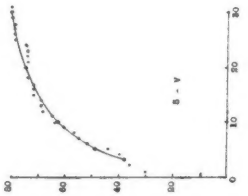

5 - V
4 - V
5 - I
4 - I
1 - V
2 - V
1 - I
2 - I

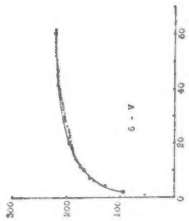

7 - V

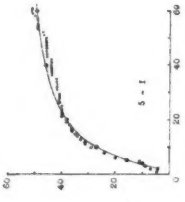

8 - V

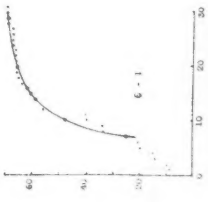

7 - I

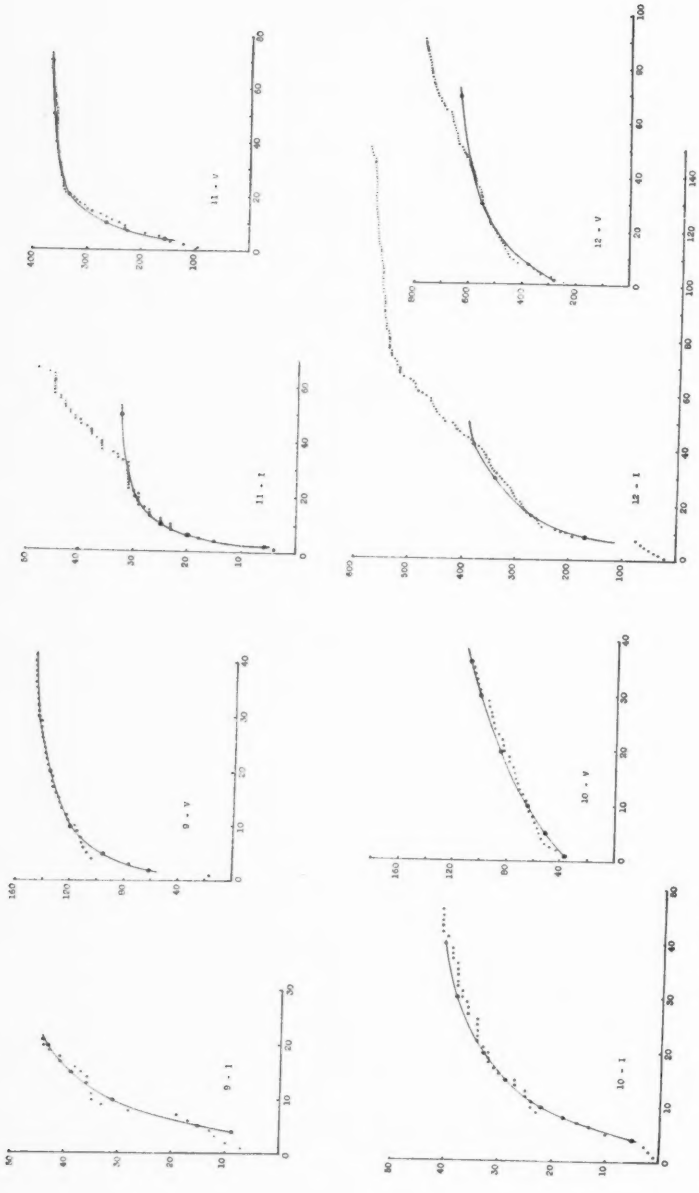

8 - I

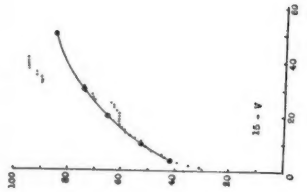

8 - V

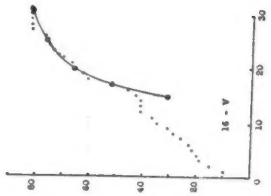

6 - V

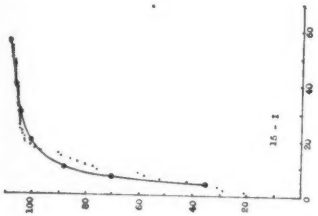

5 - I

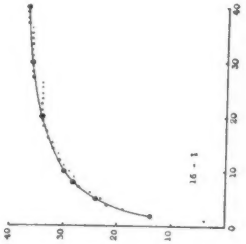

6 - I


11 - V
11 - I
12 - V
12 - I
9 - V
9 - I
10 - V
10 - I

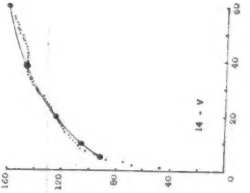

15 - V

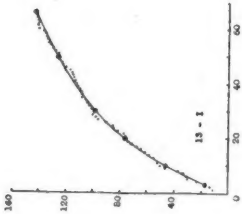

16 - V

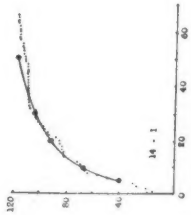

15 - I

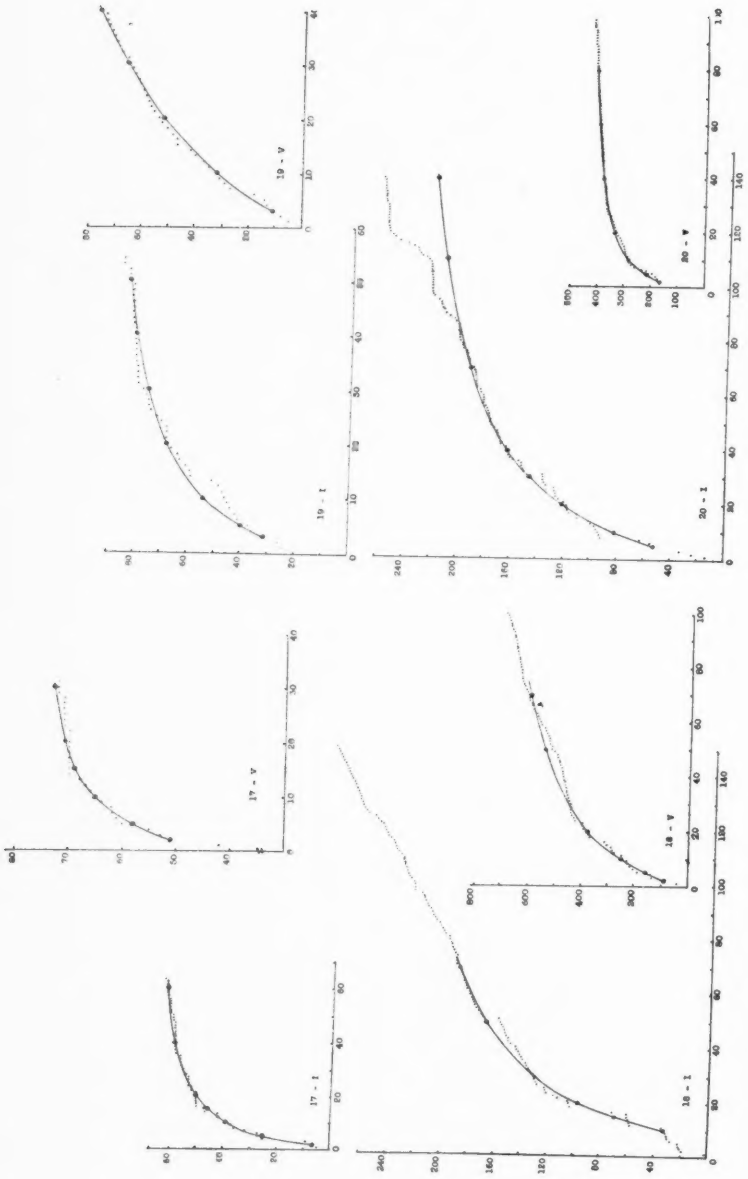

16 - I

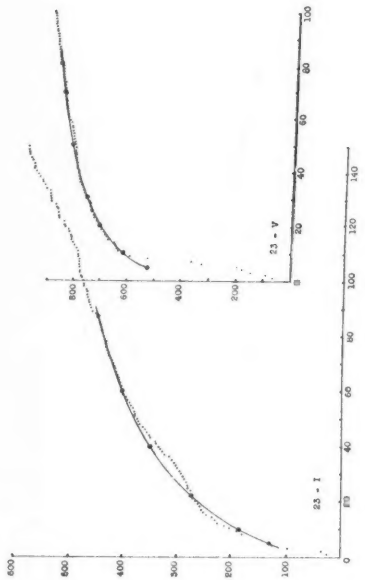
13 - V

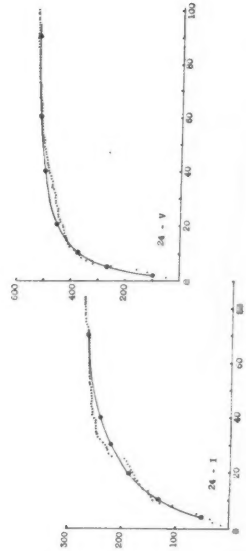

14 - V

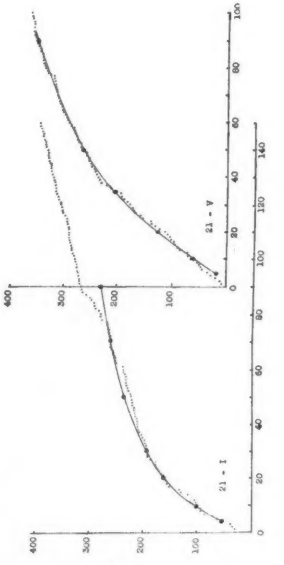

13 - I

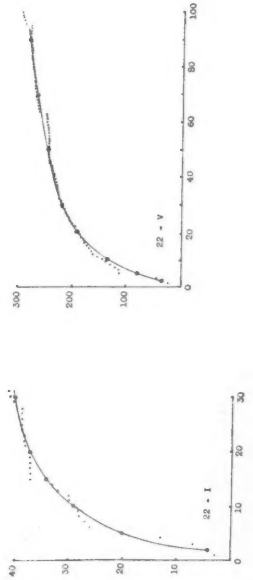

14 - I

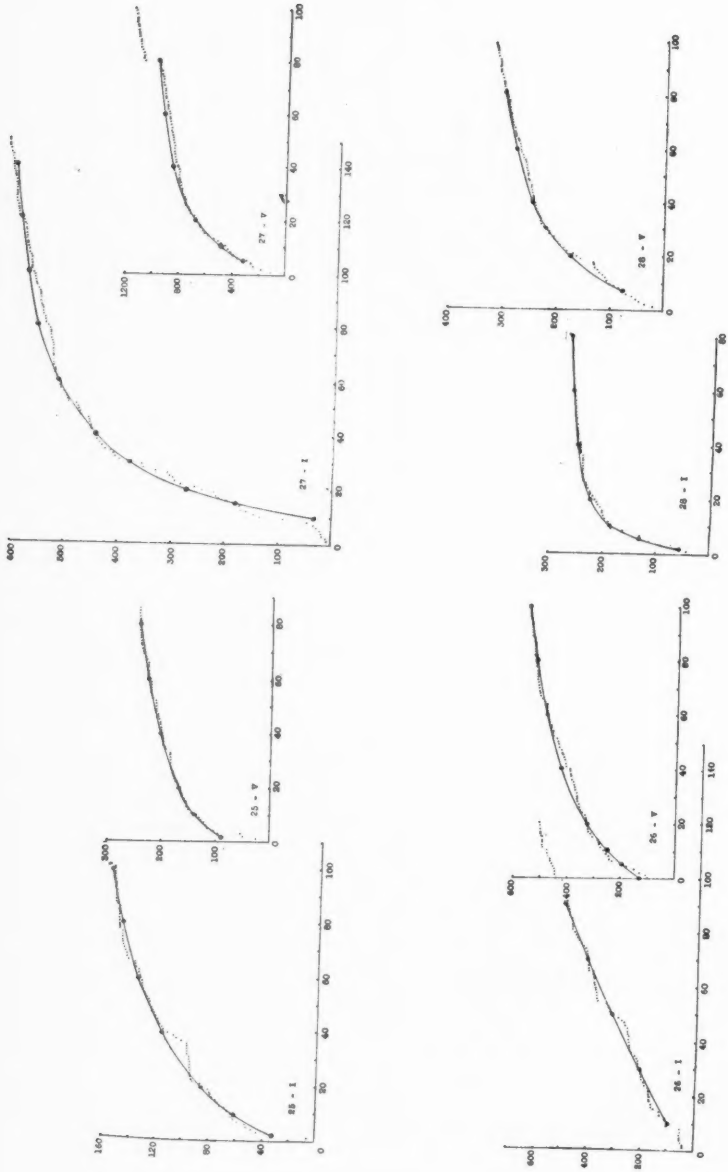

19 - V
19 - I
20 - V
20 - I
17 - V
17 - I
18 - V
18 - I

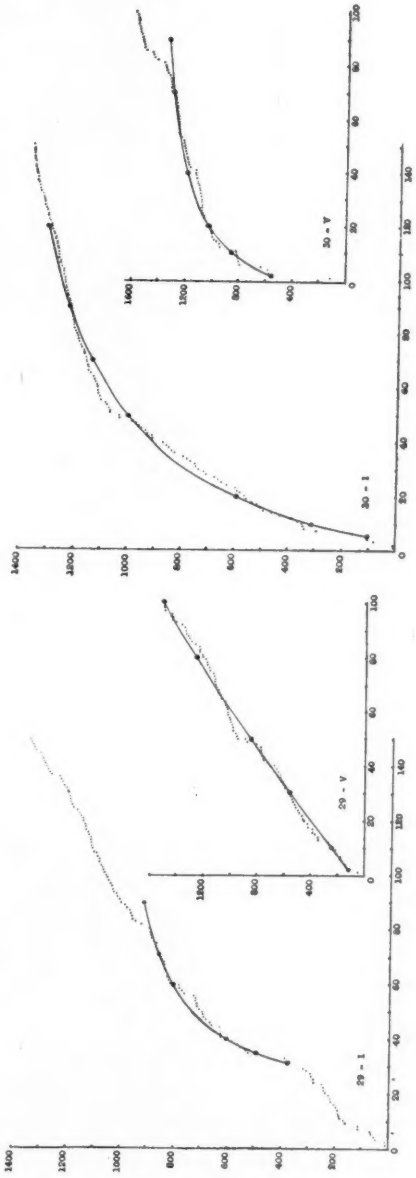

23 - V
23 - I


24 - V
24 - I


21 - V
21 - I


22 - V
22 - I


27 - I
27 - V
28 - I
28 - V
25 - I
25 - V
26 - I
26 - V


30 - V
30 - I
29 - V
29 - I

PSYCHOMETRIKA—VOL. 2, NO. 1
MARCH, 1937

MATRIX APPROXIMATIONS AND SUBSPACE FITTING

GALE YOUNG

University of Chicago*

The problem of fitting a least-squares linear subspace to a higher dimensional set of points is solved by the use of matrix approximation results, both when the subspace is required to pass through a given point, and when it is not so restricted. In the latter case the subspace is shown to pass through the centroid of the point set. It is noted that the principal axes of the set of points enter into the problem in a natural manner.

Recent work (**1**) has provided the solution of the following matrix approximation problem: given a real matrix a of rank s, to find β of rank $r < s$ which minimizes the sum of squares of elements of the difference matrix $a - \beta$†. β will be called the "best (rank r) approximation to a". This result has a geometrical interpretation which amounts to a generalization of the familar process of determining the least-squares straight line for a set of points, the generalization being from the best fitting straight line to the best fitting r-dimensional linear subspace.

Preliminary Remarks

Consider a t-dimensional euclidean space E. By a linear subspace R of r dimensions is meant the totality of points of E which are spanned by (i. e., which can be reached by taking linear combinations of) a set of r linearly independent vectors emanating from some point of E. A vector coordinate system for E is said to be adapted to R if its origin is in R and R is spanned by some r of the vectors. There is always an orthogonal coordinate system for E which is adapted to R and has any given point of R as origin.

The following lemmas are stated without proof. They will be recognized as simple extensions of propositions well known in two or three dimensional cases.

(*I*) *For any point P of E there is a unique closest point P' of R;*

*The author is indebted to Prof. Carl Eckart of the Department of Physics, University of Chicago, for much helpful comment and suggestion. To him are due the example of the last section and various other items throughout.

†A similar result obtains for the corresponding problem in complex matrices, with the above "sum of squares" replaced by "sum of squares of absolute values".

i. e., P' is at a smaller distance from P than is any other point of R.

The distance $\overline{PP'}$ is termed the distance of P from R; and P' is said to be the projection of P on R. It is in fact the perpendicular projection, $\overline{PP'}$ being one of the perpendiculars to R at P'.

Similarly, one may have an entire linear subspace T of points, and consider its projection T' on R. Then

(*II*) *T' is a linear subspace of both R and E, with number of dimensions not greater than that of T.*

An Interpretation of Matrix Approximations

Let a denote a matrix with t rows and q columns. Each column may be thought of as giving t coordinates of a point, and thus to a there corresponds a set of q points in a t-space E (or, alternatively, one might have taken the rows as points, obtaining t points in a q-space). The best approximating matrix β of rank r then specifies a set of q points lying in an r-dimensional subspace* R which passes through the origin of E. If β is unique, so is R(**1**).

(*III*) *Of all r-spaces through the origin, R so determined minimizes the sum of squares of the distances from it to the points of a. The points of β are the points of R respectively closest to the points of a.*

First, note that the best matrix approximation of rank r is better than any of lower rank (**1**). Now if there were a better fitting subspace R', the matrix β' of the points of R' respectively closest to those of a would be a better approximation (of rank $\leq r$) to a than is β. Similarly, if the β points were not respectively the closest one in R, a better matrix would be given by these.

Origin at Centroid

For the case $r = 1$, $t = 2$ the above result reduces to the least-squares straight line through the origin for a set of q points in a plane. Now just as there is in the plane case a general least-squares line which is not restricted to pass through the origin, so it is natural to inquire if this restriction can be removed in general. The fact that the best unrestricted line in the plane passes through the centroid of the point set gives the clue for the general case.

(*IV*) *If a subspace R does not pass through the centroid C of a given set of points, there is a better fitting r-dimensional one, W, which does.*

*Here, as throughout, "subspace" means "linear subspace".

First choose a unit orthogonal coordinate system for E adapted to R, and then consider a parallel system with origin at C. For W take the r-space through C which is spanned by the same vectors as is R; i. e., W is the r-space through C parallel to R. Let x_{ij} denote the coordinate of the jth point along the ith axis of the R system. The corresponding component in the W system is

$$y_{ij} = x_{ij} - c_i \ ,$$

where the c_i are the R system coordinates of C, and are not all zero for $i > r$ (the axes are supposed to be so numbered that R is spanned by the first r of them). The sums of squares of the distances of the points from W and R are

$$L_W = \sum_{j=1}^{q} \sum_{i=r+1}^{t} y^2{}_{ij} \ ;$$

and

$$L_R = \text{corresponding sum of } x^2{}_{ij} = \text{sum of } (y_{ij} + c_i)^2. \qquad (1)$$

The coordinates of the centroid of a group of points are the means of those of the points in any rectilinear coordinate system. Hence

$$\sum_j y_{ij} = 0 \ ,$$

and expression (1) reduces to

$$L_R = L_W + q \sum_{i=r-1}^{t} c^2{}_i \ .$$

Thus $L_R > L_W$, as was to be shown.

From (IV) it follows at once that any best fitting subspace must contain the centroid, and hence can be found by passing to it as origin:

(*V*) *The r-space best fitting the column points of a is found by shifting each row of a to the mean of its elements as origin, and then applying* (*III*).

It has been customary in factor analysis to so shift the rows of the score matrix (**2**). The above indicates the significance of this, and suggests that much of factor analysis may be visualized as a process of fitting a flat subspace to a higher dimensional set of points.

Best Fitting Subspaces of Different Dimensions

If the rank s of a is less than t, the points of a are contained in an s-dimensional subspace S through the origin of E. It seems reasonable to expect that S should contain the best fitting spaces of fewer

dimensions, just as the best straight line to a coplanar set of points is itself in their plane. This is in fact correct.

(VI) R of (III) is contained in S.

For if such were not the case, the projection R' on S would be a better fitting subspace. This may be seen as follows: All the columns of β may be expressed in terms of some r of them. Not all of these r points are in S, for then R would be. Hence there is at least one point of β not in S. The projection of any such point P onto S is nearer to any given point Q of S than is P itself; from which the above statement follows. Also, by (II), R' has dimensionality $\leq r$, and there is thus a contradiction to R being the best fitting r-space.

Consider now the best fitting spaces of different numbers of dimensions. It suggests itself that each of these should be contained in the next — the best line in the best plane, etc. This also is readily verified. For, denoting the best approximating matrix of rank r by β_r, it may be seen from the nature of the approximating process **(1)** that β_i is the best rank-i approximation not only to $\alpha = \beta_s$, but also to β_j for $j = i + 1, \cdots s - 1$. In view of (VI), this gives

(VII) The various dimensional subspaces through the origin which best fit a given set of points are "nested", each being contained in all with more dimensions. Further, each is the best fit to the projection of the original set of points upon any space higher in the sequence.

Principal Axes

The above sequence of subspaces, if unique, defines a unique orthogonal coordinate system which is adapted to all of them; this is in fact the system of principal axes of the point set **(3)**. If the same set of tests be given to two groups of individuals, 1 and 2, the two resulting point sets have the same principal axes if $a_1 a_2$ is symmetric, where $a_1 = a_1\ a'_1$, etc. **(1)**. In this case, when the computation has been carried through for one group, that for the other follows with almost no additional labor.*

However, when the calculations have been completed, it is the subspace R and the β points in it which are of most interest; and for the purpose of describing these one coordinate system for R is as good as any other. Thus the principal axes may be of prime importance for computation, without being primary in interpretation or

*Similar considerations would, of course, hold if two sets of tests be given to one group of individuals, but it seems less likely that the symmetry condition would be satisfied.

theory of the results obtained. As an example, in determining the plane which most nearly contains the orbit of a planet, the principal axes may vary widely from one group of observations to another, though the best plane remains nearly the same. After the plane is found, other axes are the ones of significance for celestial mechanics.

It is thus clear that any specific choice of coordinate system must depend upon additional considerations, and that different theories may lead to different choices. This phase of the problem will not be entered into here.

REFERENCES

1. CARL ECKART AND GALE YOUNG. The Approximation of One Matrix by Another of Lower Rank. *Psychometrika,* 1936, **1**, 211-218.
2. THURSTONE, L. L. *The Vectors of Mind.* Chicago: University of Chicago Press. 1935, pp. 49-50.
3. THURSTONE, L. L. *The Vectors of Mind.* Chaper 4.

PSYCHOMETRIKA—VOL. 2, NO. 1
MARCH, 1937

THE ESSENTIAL CONDITIONS OF MEASUREMENT

DOUGLAS E. SCATES

Cincinnati Public Schools

The basic conditions and characteristics of measurement are outlined. The concepts entering into the measurement must be operationally defined; there must be a satisfactory representation of the character to be measured, and a quantitative comparison between the observed degree of the character and some reference point or scale. A conception of measurement which is too narrow would unnecessarily restrict investigation in the fields of psychology and education.

This paper is an attempt to present, more or less in outline form, the conditions and characteristics of measurement which are basic. Various elements have from time to time been declared by different writers to be essential to measurement; it is the belief of the present writer, however, that such discussions have generally conceived of measurement in too narrow terms, and have accordingly postulated requirements that are unduly limiting. There are a wide variety of measurement procedures in current practice, and if one studies the large number of examples which may readily be gathered from various fields of science, he will find it difficult to maintain a narrow concept of measurement. It is our purpose here to abstract those characteristics which appear to be common to all forms of measurement, regardless of the branch of science, and set them forth briefly.*

Reduced to its simplest terms, measurement appears to require that three conditions be met. The first of these conditions is that one have *a working concept of the character to be measured.* Obvious as as this may appear, the point is nevertheless fundamental to an intelligent use of measurement. A psychologist needs a clear notion of what intelligence is before he constructs a test to measure it. Binet, for example, struggled for some years to get an idea of the character he *ought* to be trying to measure. The crystallization of a clear concept of what one desires to measure is often a critical point in the development of a field of research.

*For a more complete discussion of each of the points taken up in this paper, together with illustrations drawn from various sciences, see a paper by the present author, "How Science Measures," published in *The Journal of Experimental Education,* Vol. V, No. 3, March, 1937.

The requirement of a clear concept is pertinent not only to new fields, but it applies to well matured areas also. For example, we may consider weight, and raise the pointed question as to the conditions under which we wish the weight of an object to be constant. Should it vary at different points on the earth? At different elevations on the earth? When in air and when in a vacuum? One's procedure in measuring weight will depend upon his decision with reference to such questions. Even such a commonplace character as weight has a number of different concepts which need to be decided upon before one is ready to undertake technical work which depends on this character.

The working concept of the character to be measured cannot at the first be perfect, but one will get results which differ from what he thinks he has, or from what he hoped for, just to the extent that the concept he has carried out is not wholly valid. For good measurement, one should have a concept that is definite and detailed, but this ordinarily comes only after extensive work in a field. In early work, the concept must almost necessarily be vague and nebulous, representing a sort of groping for something that will satisfy a certain purpose. The measurements resulting from such a concept naturally contain an abundance of spurious elements. The cultivation and delimitation of a valid concept of a critical variable is in itself an important contribution to measurement in a new area.

It is of course true that one may (more or less accidentally) derive quantities which he is unable to identify; but he at least has a working concept which springs from a knowledge of his procedure, and this in itself differentiates the result from a mere random quantity. In addition, however, one should be able to describe at least some of the direct characteristics of what he has quantified, if he is to be credited with having performed a serviceable act of measurement.

We may mention three attributes which are desirable for the concept of the character to be measured. First, it should be a relatively pure abstraction. In measuring the height of a person, we are not for the time concerned with any other characteristic than height. A character may be as complex as one desires, but it must be completely disassociated from (though not necessarily statistically independent of) all elements which are not intentionally included in it. Second, the concept should be constant; that is, the character should not change from one part to another part of its range to such an extent that it must be regarded as a different character. Thus, distance remains the same throughout the entire range of radiant energy, from

the long radio waves to the short x-rays and cosmic rays—even though the other attributes of these waves change sufficiently to warrant setting up various classes for them. Third, the concept should be well defined. This point is practically forced by the first two; it has already been discussed.

The second necessary condition of measurement is *a satisfactory representation of the character to be measured,* in the amount that is exhibited by the phenomenon. By a satisfactory representation, we mean one that is perceptible, accurate, and convenient. In some cases, this is equivalent to saying that, for characteristics which are not directly perceptible, an objectifying function or agent must be found. In other cases it means that functions must be found which adapt the variable to scales or other measuring devices which are appropriate and available.

It is with respect to this second phase of measurement that a great deal of ingenuity has been exercised, with a resulting great variety of procedures. Passing by the simple cases of *direct measurement,* such as measuring an object with a ruler, we may note that science employs various types of media in order to accomplish measurement by means of *functional representation.* These may be grouped in five large classes and denoted as mathematical, physical, chemical, biological, and psychological.

The mathematical functions are represented in a simple way in such cases as measuring the distance across a pond by taking mathematical functions of observations. In more formal work this is known as triangulation and geodesy. Without giving further illustrations, we may note that mathematical functions are involved in all, or nearly all, of the instruments of observation and measurement which are used, and that they are therefore a part of all indirect measurement. For example, a micrometer caliper or any vernier scale depends upon mathematical functions.

Physical functions also enter into all or practically all of our instruments of observation and measurement in one form or another—for example, the bending of light rays in the telescope and microscope, or even the simple balance or lever which changes the direction of force. Chemical reactions are utilized in the case of titration, precipitation, and other forms of reaction. Various combinations of physical and chemical reactions are utilized in more complex forms of measurement, as when some form of energy is converted into electricity and perhaps back again into some other form.

Typical of the indirectness of much measurement in the physical fields are the measurements which can be made with the spectroscope. The heat of distant stars is indicated by variations in intensity of different wave lengths in their spectrums, and the speed with which stars are moving toward or away from the earth is indicated by displacements in spectral lines. Such measurements suggest the degree to which observations are frequently removed from the original variable which one desires to study.

While such transformations are freely made in the physical fields, in other fields of science we find biological and psychological transformations utilized. Just as there are chemical reagents, so are there biological and psychological reagents. For example, it is normal practice in the manufacture of such products as vitamins and hormones to measure their potency by physiological assay. That is, the strength of these substances is *specified* in terms of the reactions of rats, guinea pigs, cats, dogs, etc., and the strength of each batch of the product is tested through administration to some animal. For such measurements it is as appropriate to use physiological reagents as it is to use chemical and physical reactions, or changes in energy form, for measurements in their fields.

In the field of psychology, we may properly utilize psychological reagents as indicators. In certain cases we may measure physically what has already been explored and systematized by psychological reactions, such as color, pitch, tone, quality, etc. So also to a certain extent with quality of other things. But there are no physical measurements which will indicate the difficulty of various mazes; these require the services of psychological reagents. When it comes to the evaluation of art, it is scarcely possible to raise an issue as to the appropriate procedure. When further, we desire to reflect the changing interests, moods, mental sets, etc., of people, there is no recourse except to the expressed responses of human beings.

There appear to be yet other ways in which a satisfactory representation of the character to be measured may be obtained. Instead of utilizing some function of the characteristic, as expressed through one or more of the five classes of media referred to, one may resort to a measurement in terms of component elements of the character. We shall call this process *analytical representation*, and shall note four types. First, one may seek to represent a complex character by measuring some one element that he believes to be typical of it. Examples of this are, using the number of repetitions as an index of difficulty in learning nonsense syllables, or using the number of trials

as a measure of the difficulty of learning a maze. Difficulty has many components and many manifestations.

A second technique of analytical representation for purposes of measurement is to index several different elements of a general character, and use these measures individually, with no attempt to combine them in any way. For example, in the case of personality, we are usually more interested in a profile than we are in a single rating on personality in general. This is also usually true of educational achievement batteries, although here we may also be interested in the total or average.

A third procedure is to measure the strength of various components in a general character by means of statistical analysis of the general character. Available for this purpose are such techniques as multiple regression, analysis of variance, path coefficients, and factor analysis. As the fourth procedure we have just the opposite — the attempt to represent the general character through finding some proper combining function for measurements of the elements. The statistical techniques for this purpose are simple addition, addition with various weights assigned to the elements, multiple regression, and index numbers. Other special functions have been developed.

The five types of functional representation and the four types of analytical representation indicate the procedures which have been developed to deal with variables that cannot be measured directly. When one considers the extensive use of such devices he will recognize that the means of representing the character to be measured is a phase of measurement which is of the first importance. He will also recognize that measurement is not ordinarily a simple thing, such as laying a ruler beside an object. It is frequently a matter of expending great effort to find a function which is appropriate to the character and to the conditions.

This brings us to our third condition for measurement, which is *a basis of quantitative comparison.* Measurement is essentially a "more-than" or "less-than" type of comparison between a reference point (usually a mark on a scale) and the phenomenon. This comparison may in practice take a variety of forms, and is not always direct.

Reduced to its simplest terms, measurement is the comparison of any two quantities or degrees, one of which is taken as the reference point or basis of comparison for the other. For example, we have many measurements which are dichotomous, leading only to a result of more than and less than the given standard. Items on "ob-

jective" tests are usually of this type: individuals either have or do not have the ability to respond; intermediate degrees are ignored. Dichotomies are used for the purpose of measurement in many branches of science, and are exceptionally common in connection with enumeration, leading usually to proportion. For example, the relative healthiness of a city is indicated by its death rate. "Equally often noticed differences" represents a slightly different application. This usage, when analyzed, seems to be supported by sound logic.

Ranks represent other applications of dichotomies, being a series instead of a single case. That is, each case, directly or indirectly, is compared with every other case in the series, and placed accordingly, either above or below. This brings us to the point that we do not need to know the value of the standard of comparison, in terms of ordinary units on a scale, in order to measure. Ranking, for example, can in certain cases be done directly without the use of an independent scale, such as when children are being ranked in terms of height. We do not know directly or in absolute terms, the healthiness of any city used in a comparison of death rates; in fact, it is the element of comparison that gives any death rate its significance.

For practical purposes, the value of measurement is to add something to the description of a phenomenon or situation. It is to increase the significance of the description, to make it more complete and meaningful. It is entirely possible that this purpose may be better served, in some instances, through scale marks that are not equal, and which have no suggestion of units at all. The condition necessary for this to be so is for one to have vivid experiences with certain critical marks or values. When, for example, we test a child's hearing by holding a watch a given distance from his ear, we do not have any knowledge of the units of intensity of that tick, but we may nevertheless make some very useful measurements by that means. Or if we are measuring more scientifically with a 4-A audiometer, we do not need any understanding of the unit "decibel" at all; we might need to have some gross (and unequal) divisions on the scale representing "satisfactory hearing," and a few other categories having practical significance.

We find a wide variety of forms of comparison and indication in different technical fields, which do not employ scales directly, though it is possible in some instances to convert the results into ordinary units. The use of the spectroscope has been referred to; there are no units of heat nor of "miles per hour" on the spectroscope — though we find it convenient to translate careful observations into these units

by means of elaborate calculations. Readings which are not in terms of units do not necessarily lack precision.

Measurement in terms of units which are equal, fixed, and standardized, is of course ordinarily to be preferred where it is possible. In the first place, such measurements can be readily recorded and transmitted to others. In the second place, they ordinarily convey more significance to a larger number of people than do unequal units, or values having special significance because of certain unique experiences connected with them. In the third place, they facilitate — in fact they make possible — quantitative science, with its many interrelations, expressed as laws and looser functions. To utilize units which do not correspond with our number system (in the sense of equal increments) would be to inject almost hopeless confusion into problems that are at best baffling.

While recognizing these advantages, we should not make the mistake of assuming that equal units automatically carry all the significance that they might be presumed to have. One may understand a foot, but have no comprehension at all of a light-year, or of a billionth of an inch. And for our scientific purposes, we must be frank in recognizing that our equal units are conceptual artifacts, invented to tie into our particular type of numerical thinking.

As for the necessity of a true zero point, we may call attention to the fact that in the case of temperature, we do not commonly use one even though we have one. It is much more convenient for all practical purposes, and for many scientific purposes, to use a zero nearer the commonly observed quantities. We find that zero enters into mensuration in various ways, one of which is to indicate the beginning of an interval of the character to be measured. This is the case when measuring the amount of time or distance that is occupied by an event or by an object—absolute zero in time or distance being inconceivable. This same use is available to psychological or educational measurements, as when measuring growth over a period, and it affords *ratios* for the various proportions of the interval spanned, just as it does in the case of intervals of time and distance, assuming only that the units of measurement are equal. While other types of zero are serviceable, they can be forgone.

In summary, it may be said that measurement is, in essence, a fairly simple concept, requiring only that one make a quantitative comparison between some reference point and the value or degree of the character which he is working with. In practice, however, measurement takes on a great variety of forms, and often becomes ex-

tremely complex. Three conditions have been stated as essential to measurement: (1) a working concept of the character to be measured, so that one knows how to approach its measurement, and what to do with it after he has it; (2) a satisfactory means of making the degree of the character exhibited by any phenomenon perceptible, and convenient practically, so that one may make a comparison; and, (3) a quantitative comparison between the observed degree of the character and some reference point or scale.

The readily observable aspects of common measurement — such as an independent scale, fixed, equal units, and a zero point, are all desirable for certain types of measurement, but they are not essentials in the general concept; there can be many types of measurement —and fortunately so, for there are many different phenomena to be measured, and many different conditions. To attempt to restrict measurement to the confines of certain more or less arbitrary definitions is not only futile but it would be folly if it could be done. It is through being kept flexible, and through taking the many different forms in response to the needs and the conditions, that measurement remains the essential tool of science.

MARCH, 1937
PSYCHOMETRIKA—VOL. 2, NO. 1

THE RELATION BETWEEN RESULTS OBTAINABLE WITH RAW AND CORRECTED CORRELATION COEFFICIENTS IN MULTIPLE FACTOR ANALYSIS

MERRILL ROFF
Indiana University

This paper presents three theorems concerning the relation between results with obtained and corrected correlation coefficients in the Thurstone method of multiple factor analysis. (1) The rank of the correlational matrix, and thus the number of factors involved, is unaffected by correcting the obtained coefficients for attenuation. (2) The communality of a variable when the obtained coefficients have been corrected for attenuation is equal to the communality with obtained coefficients divided by the reliabilty coefficient of the variable. (3) The relationship is shown between the factorial matrix of a correlational matrix of raw correlation coefficients and the factorial matrix of a correlational matrix of corrected correlation coefficients, and a method of obtaining either of these factorial matrices from the other is indicated.

There is diversity of practice among workers in multiple factor analysis in that some analyze the obtained correlation coefficients, while others correct their coefficients for attenuation. Spearman **(5)** advises that coefficients should be corrected, and argues against some of his critics that their failure to observe a pronounced general factor is due to error in the coefficients used. However, he showed that the vanishing of his tetrads is unaffected by correcting the coefficients for attenuation. Hotelling **(3)** considers the problem briefly, and decides in favor of correlations between the estimated true scores, with the comment **(3**, p. 22), "Since we wish to deal with real quantities as far as possible, we have based our analysis into principal components upon the corrected coefficients r_{ij}."

On the other hand, Thurstone **(6)** consistently uses raw coefficients, and Kelley **(4)** similarly seems to prefer analysis of the obtained scores.

The present paper offers three theorems on the relationship between results obtained when raw and corrected coefficients are used in the correlational matrix. The results are strictly applicable only to the Thurstone factor method; to what extent they are applicable to any other possible methods of factor analysis is not discussed in the present paper.

1. *The rank r of an n by n correlational matrix is unchanged when the correlational coefficients which are its elements are corrected for attenuation.* This is a generalization to any rank of Spearman's similar theorem concerning a correlational matrix of rank one. The proof is quite simple. A given matrix A of obtained correlation coefficients can be transformed into a matrix C whose elements are the original coefficients corrected for attenuation by premultiplying and postmultiplying A by a diagonal matrix D which has as elements the values $\frac{1}{\sqrt{r_{ii}}}$, where r_{ii} denotes a reliability coefficient, and whose diagonal entries, H_i^2, are estimated in the usual way from the coefficients in C. This is illustrated in Table I, where R_{ij} denotes r_{ij} corrected for attenuation.

TABLE I

$$\begin{Vmatrix} \frac{1}{\sqrt{r_{11}}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{r_{22}}} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{\sqrt{r_{33}}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \frac{1}{\sqrt{r_{nn}}} \end{Vmatrix} \begin{Vmatrix} h_1^2 r_{12} r_{13} & \cdots & r_{1n} \\ r_{21} h_2^2 r_{23} & \cdots & r_{2n} \\ r_{31} r_{32} h_3^2 & \cdots & r_{3n} \\ \cdots & \cdots & \cdots \\ r_{n1} r_{n2} r_{n3} & \cdots & h_n^2 \end{Vmatrix} \begin{Vmatrix} \frac{1}{\sqrt{r_{11}}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{r_{22}}} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{\sqrt{r_{33}}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \frac{1}{\sqrt{r_{nn}}} \end{Vmatrix} = \begin{Vmatrix} H_1^2 R_{12} R_{13} & \cdots & R_{1n} \\ R_{21} H_2^2 R_{23} & \cdots & R_{2n} \\ R_{31} R_{32} H_3^2 & \cdots & R_{3n} \\ \cdots & \cdots & \cdots \\ R_{n1} R_{n2} R_{n3} & \cdots & H_n^2 \end{Vmatrix}$$

$$D \qquad A \qquad D \qquad = \qquad C \qquad (1)$$

The matrix D is non-singular, since the value of each of its diagonal elements exceeds unity. The rank of C is, therefore, equal to the rank of A by the theorem (1, p. 79) that "If a matrix of rank r is multiplied in either order by a non-singular matrix, the rank of the product is also r." Thus the rank of any correlational matrix is unaffected by correcting its coefficients for attenuation. This assumes, of course, that the communalities of the variables after correction are estimated from the coefficients actually to be factored. It has been suggested to me that some persons prefer to use either unity or the communality obtained for a variable before the coefficients were corrected in the place of the communalities here. For the effect of these divergent procedures, the reader may be referred to Thurstone's discussion. (**6**, pp. 98-103).

2. *When the coefficients of a correlational matrix are corrected for attenuation, the new communality of a variable is equal to the original communality divided by the reliability coefficient of the variable.* That is,

$$H_i^2 = \frac{h_i^2}{r_{ii}} ,$$

where

$H_i^2 \equiv$ the communality in the corrected matrix,
$h_i^2 \equiv$ the communality in the original matrix,
$r_{ii} \equiv$ the reliability coefficient of the variable.

The definition of a communality is given by Thurstone (**6**, p. 62). "The communality of a test is its common factor variance"; that is, its value is equal to the sum of the squares of the common factor loadings of a test. For a discussion of the properties of and reason for using this value in the diagonal cells of a matrix to be factored, the reader may turn to Thurstone's *The Vectors of Mind,* (**6**, p. 92-119). In practice, a communality is determined after factorization by summing the squares of the common factor loadings of a variable, as they appear in the factorial matrix.

If we should use in the diagonal cells of the matrix C the values obtained by premultiplying and postmultiplying the matrix A of obtained coefficients with communalities in the diagonal cells, we would have as entries $\frac{h_i^2}{r_{ii}}$. It is to be shown that $\frac{h_i^2}{r_{ii}} = H_i^2$, the communality of a variable in the matrix of corrected coefficients.

The correlational matrix A is equal to the product of its factorial matrix F and the transpose of the factorial matrix, F' (**6**, p. 70). If there are r factors involved, F will have n rows and r columns, and we may write

$$A = F F' . \qquad (2)$$

We may indicate the elements of F by a_{ij}.
Also,

$$D A D = D F F' D \qquad (3)$$

where D is the diagonal matrix defined above and used in equation (1).

If we denote the factorial matrix of C by G, and the transpose of G by G', we have also

$$C = G G' . \qquad (4)$$

The matrix G, by theorem 1 above, will also be of order n by r; its elements may be denoted b_{ij}.
But

$$C = D A D ; \tag{1}$$

so that, substituting from (3) and (4), we obtain

$$C = D F F'D = G G' . \tag{5}$$

If we multiply out the matrix products $D F$ and $F'D$, we obtain two new matrices, E and its transpose E', whose elements have the values $\frac{a_{ij}}{\sqrt{r_{ii}}}$. This can be illustrated briefly with a 3 by 2 factorial matrix and its transpose.

TABLE II

$$\begin{Vmatrix} \frac{1}{\sqrt{r_{11}}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{r_{22}}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{r_{33}}} \end{Vmatrix} \quad \begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{Vmatrix} \quad = \quad \begin{Vmatrix} \frac{a_{11}}{\sqrt{r_{11}}} & \frac{a_{12}}{\sqrt{r_{11}}} \\ \frac{a_{21}}{\sqrt{r_{22}}} & \frac{a_{22}}{\sqrt{r_{22}}} \\ \frac{a_{31}}{\sqrt{r_{33}}} & \frac{a_{32}}{\sqrt{r_{33}}} \end{Vmatrix}$$

$$D \qquad F \qquad = \qquad E \tag{6}$$

$$\begin{Vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \end{Vmatrix} \quad \begin{Vmatrix} \frac{1}{\sqrt{r_{11}}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{r_{22}}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{r_{33}}} \end{Vmatrix} \quad = \quad \begin{Vmatrix} \frac{a_{11}}{\sqrt{r_{11}}} & \frac{a_{21}}{\sqrt{r_{22}}} & \frac{a_{31}}{\sqrt{r_{33}}} \\ \frac{a_{12}}{\sqrt{r_{11}}} & \frac{a_{22}}{\sqrt{r_{22}}} & \frac{a_{32}}{\sqrt{r_{33}}} \end{Vmatrix}$$

$$F' \qquad D \qquad = \qquad E' \tag{7}$$

It can be seen that E' is the transpose of E. We thus have

$$G G' = E E' . \tag{8}$$

This is a matrix equation of the second degree which can be solved for G in terms of E, whose elements are known. The solution of this

equation is given by Cullis (**2**, pp. 334, 335) as

$$G = LE \ , \tag{9}$$

where L is any unitary orthogonal matrix of order r. That is, it is possible, with the aid of the suitable orthogonal transformation to make E equal to G. This means that if G is a factorial matrix of C, E is also a factorial matrix of C, and we have thus the equality,

$$\sum_{j=1}^{r} \frac{a_{ij}^2}{r_{ii}} = \sum_{j=1}^{r} b_{ij}^2 \ , \tag{10}$$

and

$$\frac{h_i^2}{r_{ii}} = H_i^2 \ .$$

From equation (9) an additional theorem can be directly stated.

Theorem 3. *If a factorial matrix F obtained from a matrix of raw correlation coefficients be premultiplied by a diagonal matrix which has as elements the values $\frac{1}{\sqrt{r_{ii}}}$, where the values r_{ii} are reliability coefficients of the variables, the product is a factorial matrix of the same group of variables when all their intercorrelations have been corrected for attenuation.* Similarly, it is possible to transform a factorial matrix of a correlational matrix of corrected coefficients into a factorial matrix from the uncorrected coefficients by pre-multiplying by the inverse of the diagonal matrix defined above.

REFERENCES

1. BOCHER, M. Introduction to Higher Algebra. New York: Macmillan, 1931.
2. CULLIS, C. E. Matrices and Determinoids. Vol. II. Cambridge: University Press, 1918.
3. HOTELLING, H. Analysis of a Complex of Statistical Variables into Principal Components. Baltimore: Warwick & York, 1933.
4. KELLEY, T. L. Essential traits of Mental Life. Cambridge: Harvard Univ. Press, 1935.
5. SPEARMAN, C. E. The Abilities of Man. New York: Macmillan, 1927.
6. THURSTONE, L. L. The Vectors of Mind. Chicago: Univ. of Chicago Press, 1935.

PSYCHOMETRIKA—VOL. 2, NO. 1
MARCH, 1937

THE BI-FACTOR METHOD

KARL J. HOLZINGER AND FRANCES SWINEFORD

University of Chicago

The Bi-factor Method of factor analysis is described and illustrated with a small group of fourteen tests. A detailed illustration is given of how the method may be modified to the case of overlapping group factors. It is advocated that the Bi-factor pattern in unmodified form be used to determine the adequacy of tests for the measurement of unitary traits.

I. *Introduction.* In the present paper we shall give a brief description of the Bi-factor Method of factor analysis introduced in the Preliminary Reports on the Spearman-Holzinger Unitary Trait Study.* We shall also illustrate how the method may be modified for the analysis of variables of greater complexity than assumed in the original theory.

The simplest form of the Bi-factor pattern is merely an extension of Spearman's Two-factor pattern to the case of group factors. The Spearman pattern is a theoretical frame of reference consisting of a general factor running through all variables and uncorrelated factors present in each variable. The Bi-factor pattern is also a theoretical frame of reference in which a general factor is assumed to run through all variables with specific factors in each variable, but in addition a number of uncorrelated group factors, each through two or more variables, are also included. The minimum number of factors of these three types for n variables may then be briefly summarized as follows: one general factor, n specific factors, and q group factors where q is usually much smaller than n. In the modified pattern some of the group factors may overlap.

The general plan of analysis is to re-sort all of the n tests by a combination of methods, so as to bring into small groups those tests which correlate higher amongst themselves than they do with the remaining tests.† When this is accomplished, the large table of intercorrelations will show triangles of relatively high correlations along the diagonal of the whole table. The factor weights of the general

**Preliminary Reports on Spearman-Holzinger Unitary Trait Study*, Nos. 1-8 (9 in preparation). Prepared at the Statistical Laboratory, Department of Education, University of Chicago, 1930-36.

†It is assumed that all intercorrelations are positive or insignificantly negative for mental variables. In case most of the correlations of a variable x_i with the other variables are negative, the variable $-x_i$ may be used to yield positive values.

factor are computed by selecting only one test within each sub-group (in all possible combinations). The general factor is next removed. The main body of residuals is then numerically small, but the triangles of high correlations for the group factors remain on the diagonal. The weights of these group factors may next be computed and the factors removed to furnish final residuals for study. If the latter show that no greater complexity of group factors is required, the two specific factors in each test (specificity and unreliability) may be readily calculated and the final factor pattern set up.

In case the first residuals show a complexity greater than the simple Bi-factor form, the extra factor may be introduced into what may be called the Modified Bi-factor pattern. After their proper allocation (to be illustrated below), the whole analysis is repeated in terms of the new frame of reference and new final residuals computed. These are again examined for any significant factor overlap to determine the final goodness of fit of the original correlations and modified factor pattern.

The arbitrary nature of the Bi-factor frame of reference has been commented on elsewhere as a distinct limitation of the method. By the above modification this objection would seem to disappear, but we should like to emphasize the view that without modification of the pattern, the limitation may be regarded as a defect of the tests used rather than of the method. We should go so far as to argue that no modification should be necessary if the tests are properly made to measure single group factors. If some tests do reveal two or more group factors, then they are poor tests for the purpose of factor appraisal, and should be improved or discarded. In short, the Bi-factor frame of reference may serve as a guide to the construction of tests as measures of factor ability, as well as a very simple and easy basis for analysis.

II. *The Bi-factor Pattern.* In the present analysis it is assumed that all variables are represented in standard form and that all factors are uncorrelated. The chief advantages arising from assuming uncorrelated factors are in the simplicity of analysis and economy of measurement. If two factors are uncorrelated, a measure of one does not involve the other, a difficulty which has made present-day testing confusing and well-nigh hopeless. Thus an ordinary reading comprehension test and a verbal intelligence test will be highly correlated, and these two labels almost useless as indexes of these traits considered as two abilities. For economical measurement, simplicity, and parsimony, uncorrelated factors are indispensable.

The present notation is as follows:

$N \equiv$ number of individuals in the sample
$n \equiv$ total number of tests
$x_i \equiv$ standard score for test i
$r_{ij} \equiv$ product-moment correlation between tests i and j.
$a \equiv$ general factor (designated as u_1 in the Reports* and in the numerical example)
β, γ, δ, etc. $\equiv$ group factors
$a_i \equiv$ weight of general factor for test i
b_i, c_i, d_i, etc. $\equiv$ weights for group factors β, γ, δ, etc.
s_i and $t_i \equiv$ weights of the specific and unreliability factors, respectively
$\underset{\sim}{r}_{ij} = r_{ij} - a_i a_j \equiv$ a residual correlation with factor a removed from tests i and j.
$r_{11} \equiv$ reliability coefficient

A hypothetical factor pattern written in tabular form is as follows:

TABLE I

HYPOTHETICAL BI-FACTOR PATTERN

Variable	α	β	γ	δ	Specificity s_i	Unreliability $t_i = \sqrt{1-r_{11}}$
1	a_1	b_1			s_1	t_1
2	a_2	b_2			s_2	t_2
3	a_3	b_3			s_3	t_3
4	a_4	b_4			s_4	t_4
5	a_5		c_5		s_5	t_5
6	a_6		c_6		s_6	t_6
7	a_7		c_7		s_7	t_7
8	a_8		c_8		s_8	t_8
9	a_9			d_9	s_9	t_9
10	a_{10}			d_{10}	s_{10}	t_{10}
11	a_{11}			d_{11}	s_{11}	t_{11}
12	a_{12}			d_{12}	s_{12}	t_{12}
	Σa^2_i	Σb^2_i	Σc^2_i	Σd^2_i	Σs^2_i	Σt^2_i

The weights for the specificity and unreliability factors should be staggered in the pattern, so that no two s's or t's appear in the same column, but they are written here in more compact form to save space.

All of the weights represent the loadings of the test with the fac-

op. cit.

tor in question, or the correlation of the factor with the variable they resolve. Thus b_3 is the correlation of the β factor with test 3, etc.

From the above assumptions, the following relations due to Spearman may be noted:

$$r_{ij} = a_i a_j + b_i b_j \ , \tag{1}$$

$$1 = a_i^2 + b_i^2 + s_i^2 + t_i^2 \ , \tag{2}$$

$$t_i^2 = 1 - r_{1I} \text{ where } r_{1I} = \text{reliability} \ , \tag{3}$$

$$n = \text{sum of squares of all factor weights} \ , \tag{4}$$

$$\sum_1^n a_i^2 = \text{portion of total variance due to } a \ . \tag{5}$$

In the original formulation of the Bi-factor theory, weights such as b_1, b_2, and b_3 were taken proportional to the corresponding weights a_1, a_2, and a_3, but this is not necessary for analysis as pointed out in *Preliminary Report* No. 7.

If we assume that the factors are allocated as in the hypothetical example above, the a factor may next be removed by taking only one test from each sub-group in all possible combinations to compute the necessary weights. Thus to compute a_1 we use tests 1, 5, and 9; 1, 6, and 9; etc., to give,

$$r_{1a}^2 = a_1^2 = \frac{r_{15}r_{19} + r_{16}r_{19} + \text{etc.}}{r_{59} + r_{69} + \text{etc.}} \ . \tag{6}$$

A complete outline for performing these calculations from simple sums has been described in *Preliminary Report* No. 7.

When the weights a_i are thus found the products $a_i a_j$ are formed and then the residuals $r_{ij} = r_{ij} - a_i a_j$ determined. The group factors may next be removed in any order by the direct use of Spearman's formula

$$r_{i\beta}^2 = b_i^2 = \frac{\left[\sum_1^{m-1} r_{c_x}^2 - \sum_1^{m-1} r_{c_x}^2\right]}{2\left[\sum_1^{C_2^m} r - \sum_1^{m-1} r_{c_x}\right]}, \tag{7}$$

where m is the number of tests in the sub-group (here four), r_{c_x} the correlation in a column of these correlations for the group factor, and r a correlation in the set of C_2^m values (here six).

The weights t_i are obtained directly from the reliability coefficients r_{1I}. Thus $t_1 = \sqrt{1 - r_{1I}}$ where r_{1I} = reliability of test 1. The values for s_i are found from the relation $s_i^2 = 1 - a_i^2 - b_i^2 - t_i^2$.

If the residuals with a removed reveal any extra overlap, it is desirable to revise the pattern plan before computing the group factors. A frequency distribution of final residuals is made for the modified pattern and the value .6745 σ of this distribution compared with the probable error of a zero correlation for a sample of N. This will serve as an approximate standard for judging the goodness of fit of the Bi-factor pattern to the observed correlations. If the value .6745 σ of the residuals is appreciably less than this standard it is probable that insignificant group factors have been introduced into the pattern.

III. *Allocation of Tests to Groups.* We may now turn to the general problem of analyzing a set of observed variables by the Bi-factor method. In allocating the tests for the determination of group factors, three procedures have been followed:

(a) The use of a B-coefficient which is merely the average of all intercorrelations of tests 1, 2, $\cdots$, m, divided by the average of all correlations of tests 1, 2, $\cdots$, m with the remaining tests not in this group. The quotient expresses the extent to which the tests 1, 2, $\cdots$, m belong together in the sense that they have high correlations amongst themselves and relatively low correlations with other tests in the whole set.

The work is begun by selecting the two tests with the highest correlation in the whole large table, and then the test which correlates most highly with either of these, and proceeding likewise until a drop in B value is obtained.

(b) The nature of the tests themselves is carefully studied as the B-coefficients are computed. If the coefficients fail to show appreciable drop, tests are sometimes included or rejected upon examination of their content.

(c) The residuals are always reproduced in full to reveal overlap, extra factors, or the wrong allocations of tests. If a test has been omitted from a group to which it belongs, consistently positive residuals will occur between this test and those of the group. If, on the other hand, a test has been included which does not belong in the group, a negligible loading for the group factor usually is found.

These three methods have been used successfully on several sets of correlations described in our recent *Preliminary Reports.*

TABLE II

Partial Correlations with Age Eliminated

(Fourteen Tests; 355 Cases)

Test	1	2	3-4	6	28	29	32	34	35	36a	13	18	25b	77
1														
2	.514													
3-4	.477	.662												
6	.433	.497	.415											
28	.424	.397	.319	.444										
29	.350	.427	.376	.530	.437									
32	.083	.152	.173	.064	.027	.018								
34	.239	.254	.172	.371	.211	.224	.264							
35	.140	.083	.137	.214	.139	.066	.203	.334						
36a	.286	.368	.229	.394	.267	.340	.191	.442	.234					
13	.305	.545	.482	.354	.262	.349	.166	.202	.007	.360				
18	.260	.526	.373	.348	.193	.358	.115	.159	—.014	.372	.677			
25b	.231	.437	.424	.310	.160	.245	.129	.058	—.030	.235	.603	.596		
77	.250	.426	.368	.279	.189	.273	.133	.039	—.037	.241	.586	.613	.559	

TABLE III

B-Coefficients

B(13, 18)	2.00
B(13, 18, 77)	2.14
B(13, 18, 77, 25b)	2.40
B(13, 18, 77, 25b, 2)	2.18
B(13, 18, 77, 25b, 2, 3-4)	2.20
B(2, 3-4)	1.85
B(2, 3-4, 1)	1.72
B(2, 3-4, 1, 6)	1.59
B(2, 3-4, 1, 6, 29)	1.60
B(6, 29)	1.68
B(6, 29, 28)	1.67
B(6, 29, 28, 36a)	1.43
B(6, 29, 28, 36a, 34)	1.40
B(34, 36a)	1.75
B(34, 36a, 35)	1.75
B(34, 36a, 35, 32)	1.64
B(34, 36a, 35, 32, 6)	1.30
B(34, 36a, 35, 32, 6, 29)	1.13

IV. *A Bi-factor Analysis of Fourteen Tests.* Our most recent analysis, now completed, includes the factorization of forty-four tests for 355 cases. We have used this material to select a small group of fourteen tests to illustrate the modified Bi-factor pattern. This subgroup of tests contains the only portion of the total battery presenting necessity for modification. Brief descriptions of the fourteen tests are given in *Preliminary Reports* 1 and 3.

The set of ninety-one intercorrelations is presented as Table II. The tests have been grouped, for convenience, in accordance with the interpretation of the B-coefficients of Table III.*

The *B*-coefficients suggest the following pattern, in which x represents a factor weight different from zero (specific factors not indicated):

TABLE IV

FIRST PATTERN PLAN

Test	Factors				
	u General	*e* Spatial	*a* Mental Speed	*m* Motor Speed	*v* Verbal
1	x	x			
2	x	x			
3-4	x	x			
6	x		x		
28	x		x		
29	x		x		
32	x			x	
34	x			x	
35	x			x	
36a	x			x	
13	x				x
18	x				x
25b	x				x
77	x				x

The values of r_{iu} for this pattern are next found by means of formula (6)†, and are recorded in Table V. The ninety-one products $r_{iu}r_{ju}$ are subtracted from the corresponding entries of Table II to give the residual correlations of Table VI. If the pattern plan is a reasonable one, the values in bold-face type will be used for the calculation of the group factors and the remaining values (in roman type) should be negligible.

*Computed by method of *Report 7*, pp. 3-5.
†See *Report* 7, pp. 8-10, for detailed outline of procedure.

TABLE V

VALUES OF r_{iu}

Test i	r_{iu}
1	.615+
2	.828
3-4	.697
6	.703
28	.489
29	.570
32	.168
34	.323
35	.090
36a	.522
13	.599
18	.532
25b	.424
77	.417

TABLE VI

RESIDUAL CORRELATIONS WITH u ELIMINATED

Test	1	2	3-4	6	28	29	32	34	35	36a	13	18	25b	77
1														
2	**.005**													
3-4	**.048**	**.085**												
6	.001	—.085	—.075											
28	.123	—.008	—.022	**.100**										
29	—.001	—.045	—.021	**.129**	**.158**									
32	—.020	.013	.056	—.054	—.055	—.078								
34	.040	—.013	—.053	.144	.053	.040	**.210**							
35	.085	.008	.074	.151	.095	.015	**.188**	**.305**						
36a	—.035	—.064	—.135	.027	.012	.042	**.103**	**.273**	**.187**					
13	—.063	.049	.064	—.067	—.031	.008	.065	.009	—.047	.047				
18	—.067	.086	.002	—.026	—.067	.055	.026	—.013	—.062	.094	**.358**			
25b	—.030	.086	.128	.012	—.047	.003	.058	—.079	—.068	.014	**.349**	**.370**		
77	—.006	.081	.077	—.014	—.015	.035	.063	—.096	—.075	.023	**.336**	**.391**	**.382**	

Examination of the residual correlations reveals a tendency toward an overlap between the mental-speed tests and the motor-speed tests with the exception of Test 32. Accordingly, a second pattern plan (Table VII) is set up. The new plan differs from the original one only through the addition of three factor weights, namely, $r_{(34)a}$, $r_{(35)a}$, and $r_{(36a)\ a}$.

TABLE VII

SECOND PATTERN PLAN

Test	Factor				
	u'	e	a	m	v
1	x	x			
2	x	x			
3-4	x	x			
6	x		x		
28	x		x		
29	x		x		
32	x			x	
34	x		x	x	
35	x		x	x	
36a	x		x	x	
13	x				x
18	x				x
25b	x				x
77	x				x

In order to compute the revised values, it must be remembered that in the Bi-factor method each r_{iu} is evaluated from triplets of tests no two of which are assumed to contain the same group factor. While the Bi-factor method is one which most readily lends itself to routine calculation, the same principle of selecting tests in threes having no common group factor may be applied to more complicated types of pattern. In the present example, it now becomes necessary

TABLE VIII

VALUES OF $r_{iu'}$

Test i	$r_{iu'}$
1	.621
2	.891
3-4	.751
6	.584
28	.417
29	.511
32	.168
34	.257
35	.076i
36a	.480
13	.632
18	.557
25b	.463
77	.457

to eliminate from the original formulae for r_{iu} each set of three tests which contains any of the following pairs:

6 - 34	28 - 34	29 - 34
6 - 35	28 - 35	29 - 35
6 - 36a	28 - 36a	29 - 36a

The adjustment may readily be made on the work-sheet which was used to calculate the coefficients, r_{iu}. The new coefficients, $r_{iu'}$ are given in Table VIII. Products and residual correlations are computed as before, the latter appearing in Table IX.

TABLE IX

RESIDUAL CORRELATIONS WITH u' ELIMINATED

Test	1	2	3-4	6	28	29	32	34	35	36a	13	18	25b	77
1														
2	**—.039**													
3-4	**.011**	**—.007**												
6	.070	—.023	—.024											
28	.165	.025	.006	**.200**										
29	.033	—.028	—.008	**.232**	**.224**									
32	—.021	.002	.047	—.034	—.043	—.068								
34	.079	.025	—.021	**.221**	**.104**	**.093**	**.221**							
35	.140	.083	.137	**.214**	**.139**	**.066**	**.203**	**.334**						
36a	—.012	—.060	—.131	**.114**	**.067**	**.095**	**.110**	**.319**	**.234**					
13	—.087	—.018	.007	—.015	—.002	.026	.060	.040	.007	.057				
18	—.086	.030	—.045	.023	—.039	.073	.021	.016	—.014	.105	**.325**			
25b	—.057	.024	.076	.040	—.033	.008	.051	—.061	—.030	.013	**.310**	**.338**		
77	—.034	.019	.025	.012	—.002	.039	.056	—.078	—.037	.022	**.297**	**.358**	**.347**	

In Table IX the two largest positive residual correlations among those expected to approximate zero are $\underline{r}_{(1)(28)}$ and $\underline{r}_{(1)(35)}$. It is not unlikely that Test 1, a fairly easy, timed test, should measure the mental-speed factor, *a*. Table IX indicates further that the factor *e* probably is not significant for this pattern. Two of the three residual correlations assumed to contain this factor are negative. These two revisions are made in the third pattern plan (Table X). It should be noted in passing that all the revising was done at once in the larger battery of forty-four tests. No third plan was necessary. The fact that not all the overlapping was apparent in Table VI may be attributed to the unreliability of pattern weights which are estimated from only a small number of tests.

TABLE X

THIRD PATTERN PLAN

Test	Factor			
	u''	a	m	v
1	x	x		
2	x			
3-4	x			
6	x	x		
28	x	x		
29	x	x		
32	x		x	
34	x	x	x	
35	x	x	x	
36a	x	x	x	
13	x			x
18	x			x
25b	x			x
77	x			x

The Bi-factor method, together with certain modifications, is again employed for the computation of the new general factor weights, $r_{iu''}$. A work-sheet corresponding to Tables 4 and 5 on page 18 of *Report 7* may be used, with the tests arranged in the following four groups.

	Tests
1.	2
2.	3-4
3.	$a + m =$ 1, 6, 28, 29, 32, 34, 35, 36a
4.	v = 13, 18, 25b, 77

Since $r_{(32)a}$ is assumed to be zero, a number of additional sets of three tests are available:

1 — 32 — 2	28 — 32 — 2
1 — 32 — 3-4	28 — 32 — 3-4
1 — 32 — v	28 — 32 — v
6 — 32 — 2	29 — 32 — 2
6 — 32 — 3-4	29 — 32 — 3-4
6 — 32 — v	29 — 32 — v

The resulting coefficients appear in Table XI. Products and residual correlations are computed for the third time. In table XII there are no further overlaps or adjustments apparent.

TABLE XI

VALUES OF $r_{iu''}$

Test i	$r_{iu''}$
1	.533
2	.896
3-4	.768
6	.546
28	.381
29	.481
32	.175+
34	.237
35	.033i
36a	.436
13	.658
18	.576
25b	.500
77	.485—

TABLE XII

RESIDUAL CORRELATIONS WITH u'' ELIMINATED

Test	1	2	3-4	6	28	29	32	34	35	36a	13	18	25b	77
1														
2	.036													
3-4	.068	—.026												
6	**.142**	.008	—.004											
28	**.221**	.056	.026	**.236**										
29	**.094**	—.004	.007	**.267**	**.254**									
32	—.010	—.005	.039	—.032	—.040	—.066								
34	**.113**	.042	—.010	**.242**	**.121**	**.110**	**.223**							
35	**.140**	.083	.137	**.214**	**.139**	**.066**	**.203**	**.334**						
36a	**.054**	—.023	—.106	**.156**	**.101**	**.130**	.115	**.339**	**.234**					
13	—.046	—.045	—.023	—.005	.011	.033	.051	.046	.007	.073				
18	—.047	.010	—.069	.034	—.026	.081	.014	.022	—.014	.121	**.298**			
25b	—.035	—.011	.040	.037	—.030	.005	.041	—.060	—.030	.017	**.274**	**.308**		
77	—.009	—.009	—.004	.014	.004	.040	.048	—.076	—.037	.030	**.267**	**.334**	**.317**	

The third pattern plan having been accepted as final, the group factor loadings are computed from the bold-face residual correlations of Table XII. In the case of the a factor, care must be taken not to include any set of three tests which includes two of Tests 34, 35, and 36a. The m factor weights are computed from the residual correlations with u'' and a eliminated. The v factor weights are computed from the residual correlations of Table XII. The final group factor residual correlations are given in Table XIII.

TABLE XIII

RESIDUAL CORRELATIONS WITH u'' AND GROUP FACTORS ELIMINATED

Tests	1	6	28	29	32	34	35	36a
1								
6	—.044							
28	.063	—.055						
29	—.034	.031	.054					
32								
34	.012	.056	—.037	—.018	—.004			
35	.044	.036	—.011	—.056	.035	—.031		
36a	—.022	.015	—.018	.033	—.024	.040	—.004	

Tests	13	18	25b	77
13				
18	.009			
25b	.004	—.013		
77	—.012	.003	.008	

Table XIV is the completed factor pattern, describing the fourteen tests in terms of one general factor, three group factors, and specifics. The general factor turns out here to be the function meas-

TABLE XIV

FACTOR PATTERN

Test	Factor					
	u''	a	m	v	s_i	t_i
1	.533	.318			.710	.332
2	.896				.265i	.517
3-4	.768				.530	.360
6	.546	.586			.461	.382
28	.381	.496			.695+	.354
29	.481	.403			.738	.249
32	.175+		.377		.794	.444
34	.237	.318	.603		.570	.393
35	.033i	.303	.446		.709	.455+
36a	.436	.240	.370		.660	.424
13	.658			.494	.411	.393
18	.576			.586	.296	.487
25b	.500			.547	.407	.534
77	.485—			.564	.548	.382
Total Variance	3.877	1.103	.842	1.205—	4.571	2.401
Per cent Variance	27.69	7.88	6.01	8.61	32.65	17.15+

ured by Professor Spearman's Visual Perception Tests, two of which measure none of the group factors by this analysis.

The frequency distribution of the ninety-one final residual correlations is presented in the last column of Table XV in order that the value .6745 σ might be compared with the probable error of a zero correlation for 355 cases, as noted in Section II. The roman entries of Tables VI and IX are also included to show the extent to which proper adjustment of the pattern plan reduces the variability of the residual correlations. The factorization of the present small set of tests may have been over-refined in our effort to illustrate methods of modifying the Bi-factor pattern, as .6745 σ = .0285 $<$.0358.

TABLE XV

FREQUENCY DISTRIBUTIONS OF FINAL RESIDUAL CORRELATIONS

Value	Tables				
	VI	IX	XII	XIII	XII + XIII
.150 - .169	1	1			
.130 - .149	1	2	1		1
.110 - .129	2	....	1		1
.090 - .109	2	1	....		....
.070 - .089	6	5	3		3
.050 - .069	7	4	3	3	6
.030 - .049	6	6	12	6	18
.010 - .029	8	12	7	2	9
—.010 - .009	9	8	14	6	20
—.030 - —.011	9	10	8	7	15
—.050 - —.031	5	7	7	4	11
—.070 - —.051	10	4	3	2	5
—.090 - —.071	5	3	1		1
—.110 - —.091	1	....	1		1
—.130 - —.111	....	....			
—.150 - —.131	1	1			
Total	73	64	61	30	91
Mean	.0053	.0082	.0077	.0008	.0054
S.D.	.0616	.0558	.0450	.0354	.0422
.6745(S.D.)	.0416	.0376	.0304	.0239	.0285
PE of zero r	.0358	.0358	.0358	.0358	.0358

The Bi-factor analysis illustrated above is not only very simple, but the calculation is relatively easy as compared with other methods. The total time for computation, done by one person, was less than ten hours for the present example.

PSYCHOMETRIKA—VOL. 2, NO. 1
MARCH, 1937

A SEVEN-DECIMAL TABLE OF THE AREA (α) UNDER THE UNIT NORMAL CURVE, FOR ABSCISSAE EXPRESSED IN TERMS OF *P.E.*

RUTH H. KRAUSE AND HERBERT S. CONRAD*

University of California

A seven-decimal table is presented of the area under the unit normal curve, for abscissae expressed in terms of the "probable error" or *PE*. From the method of calculation, the partial verification by means of other tables, and the safeguards taken in the routine of computation, it is safe to consider the table reliable. Errors in previously published tables are pointed out.

Sound statistical practice requires that every statistical constant should, when possible, be accompanied by a measure of its probable fluctuation due to random errors. The measure of random fluctuation most commonly used is the "probable error" or *PE*. Thus, one customarily reports the mean and its *PE*, the standard deviation and its *PE*, etc. The practical importance of the concept of "probable error" is attested by the inclusion of this measure in virtually all current textbooks on statistics. These facts suggest the desirability of an accurate table of the area under the unit normal curve, for abscissae expressed in terms of *PE*.†

Table 1 gives the area (α) under the unit normal curve between

*For criticism of the manuscript and valuable suggestions, the writers are indebted to Professors Raymond T. Birge, Albert H. Mowbray, Noel Keys, and Miss Louise Boulware. For assistance in certain portions of the computations, we wish to acknowledge the services of Mrs. Mildred E. Conrad and Mrs. Lina H. Aylesworth.

†The standard table for the area under the unit normal curve, for abscissae expressed in units of σ or standard error, has been prepared by W. F. Sheppard (**27**, **31**). It is, of course, possible to convert *PE*-units to σ-units (from the relation, $1\ PE = .6744897502\,\sigma$); and then to find the area corresponding to a given *PE*-deviation by interpolation into Sheppard's table. The necessary multiplication by .6744897502, however, constitutes an extra step; and the task of accurate interpolation is a further additional step, not without its computational hazard. Finally, it should be remembered that an interpolated value of the area corresponding to a given *PE*-deviation will not, in general, possess the same degree of accuracy as the original values in the table from which it was obtained.

TABLE 1

TABLE OF THE AREA (a) UNDER THE UNIT NORMAL CURVE, FOR ABSCISSAE EXPRESSED IN TERMS OF *P.E.*

x / *P.E.*	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09	x / *P.E.*
.00	.0000000	.0053816	.0107630	.0161438	.0215240	.0269031	.0322811	.0376576	.0430323	.0484051	.00
.10	.0537757	.0591439	.0645094	.0698719	.0752313	.0805872	.0859395	.0912880	.0966322	.1019721	.10
.20	.1073074	.1126379	.1179632	.1232833	.1285977	.1339064	.1392090	.1445054	.1497952	.1550783	.20
.30	.1603545–	.1656234	.1708850–	.1761389	.1813849	.1866227	.1918523	.1970733	.2022855	.2074887	.30
.40	.2126827	.2178673	.2230422	.2282072	.2333621	.2385067	.2436408	.2487641	.2538765	.2589778	.40
.50	.2640677	.2691460	.2742126	.2792672	.2843096	.2893396	.2943571	.2993618	.3043536	.3093322	.50
.60	.3142974	.3192492	.3241871	.3291112	.3340212	.3389169	.3437982	.3486648	.3535167	.3583535–	.60
.70	.3631752	.3679815–	.3727723	.3775475	.3823069	.3870502	.3917774	.3964883	.4011827	.4058604	.70
.80	.4105214	.4151655–	.4197925–	.4244022	.4289946	.4335694	.4381266	.4426660	.4471875	.4516909	.80
.90	.4561764	.4606430	.4650915–	.4695213	.4739325–	.4783248	.4826982	.4870525	.4913877	.4957035	.90
1.00	.5000000	.5042770	.5085343	.5127720	.5169898	.5211878	.5253657	.5295235	.5336612	.5377785	1.00
1.10	.5418755	.5459521	.5500081	.5540435–	.5580582	.5620521	.5660252	.5699774	.5739086	.5778188	1.10
1.20	.5817078	.5855757	.5894223	.5932477	.5970517	.6008343	.6045955–	.6083351	.6120532	.6157497	1.20
1.30	.6194246	.6230779	.6267094	.6303192	.6339072	.6374734	.6410177	.6445402	.6480408	.6515195	1.30
1.40	.6549763	.6584111	.6618240	.6652149	.6685838	.6719307	.6752556	.6785585–	.6818393	.6850982	1.40
1.50	.6883350	.6915499	.6947427	.6979135	.7010623	.7041892	.7072940	.7103770	.7134379	.7164770	1.50
1.60	.7194941	.7224894	.7254628	.7284143	.7313441	.7342521	.7371383	.7400029	.7428457	.7456669	1.60
1.70	.7484665–	.7512445	.7540010	.7567360	.7594496	.7621418	.7648126	.7674621	.7700904	.7726975–	1.70
1.80	.7752834	.7778483	.7803921	.7829149	.7854168	.7878979	.7903582	.7927977	.7952166	.7976149	1.80
1.90	.7999926	.8023499	.8046868	.8070033	.8092997	.8115758	.8138318	.8160678	.8182839	.8204801	1.90
2.00	.8226564	.8248131	.8269502	.8290677	.8311657	.8332444	.8353037	.8373438	.8393649	.8413669	2.00
2.10	.8433499	.8453141	.8472595	.8491863	.8510944	.8529841	.8548554	.8567084	.8585432	.8603599	2.10
2.20	.8621586	.8639393	.8657023	.8674475	.8691751	.8708852	.8725779	.8742533	.8759114	.8775525–	2.20
2.30	.8791765	.8807837	.8823740	.8839476	.8855047	.8870452	.8885694	.8900773	.8915690	.8930447	2.30
2.40	.8945044	.8959483	.8973764	.8987888	.9001858	.9015673	.9029335	.9042845	.9056204	.9069414	2.40
2.50	.9082474	.9095387	.9108153	.9120774	.9133250–	.9145583	.9157773	.9169823	.9181732	.9193503	2.50
2.60	.9205135	.9216631	.9227991	.9239216	.9250308	.9261268	.9272096	.9282794	.9293363	.9303804	2.60
2.70	.9314117	.9324305	.9334368	.9344307	.9354124	.9363819	.9373393	.9382848	.9392185–	.9401404	2.70
2.80	.9410507	.9419495–	.9428368	.9437129	.9445777	.9454315–	.9462742	.9471060	.9479271	.9487375–	2.80
2.90	.9495372	.9503265	.9511055–	.9518741	.9526326	.9533809	.9541193	.9548479	.9555666	.9562757	2.90
3.00	.9569752	.9576652	.9583459	.9590172	.9596793	.9603324	.9609765–	.9616116	.9622380	.9628556	3.00
3.10	.9634646	.9640651	.9646571	.9652408	.9658163	.9663836	.9669428	.9674940	.9680373	.9685729	3.10
3.20	.9691007	.9696209	.9701335	.9706387	.9711366	.9716271	.9721105–	.9725867	.9730559	.9735181	3.20
3.30	.9739735	.9744221	.9748640	.9752993	.9757280	.9761502	.9765661	.9769757	.9773790	.9777762	3.30
3.40	.9781673	.9785523	.9789315	.9793048	.9796724	.9800342	.9803904	.9807410	.9810861	.9814259	3.40

	.9793048	.9796724	.9800342	.9803904	.9807410	.9810861	.9814259	3.40

TABLE 1 (continued)

TABLE OF THE AREA (a) UNDER THE UNIT NORMAL CURVE, FOR ABSCISSAE EXPRESSED IN TERMS OF *P.E.*

x / *P.E.*	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09	x / *P.E.*
3.50	.9817602	.9820893	.9824132	.9827319	.9830455	.9833542	.9836579	.9839567	.9842507	.9845399	3.50
3.60	.9848245–	.9851044	.9853798	.9856507	.9859171	.9861792	.9864370	.9866905–	.9869398	.9871849	3.60
3.70	.9874260	.9876631	.9878961	.9881253	.9883506	.9885722	.9887899	.9890041	.9892145–	.9894213	3.70
3.80	.9896246	.9898245–	.9900208	.9902139	.9904035	.9905899	.9907731	.9909531	.9911299	.9913036	3.80
3.90	.9914743	.9916420	.9918067	.9919685	.9921275–	.9922836	.9924370	.9925876	.9927355–	.9928808	3.90
4.00	.9930234	.9931635–	.9933010	.9934361	.9935687	.9936988	.9938266	.9939521	.9940753	.9941962	4.00
4.10	.9943148	.9944313	.9945456	.9946578	.9947679	.9948760	.9949820	.9950861	.9951882	.9952883	4.10
4.20	.9953866	.9954830	.9955776	.9956704	.9957614	.9958507	.9959383	.9960241	.9961084	.9961910	4.20
4.30	.9962720	.9963515–	.9964294	.9965058	.9965807	.9966541	.9967261	.9967967	.9968659	.9969337	4.30
4.40	.9970002	.9970653	.9971292	.9971918	.9972532	.9973133	.9973722	.9974299	.9974865	.9975420	4.40
4.50	.9975963	.9976495–	.9977016	.9977527	.9978027	.9978517	.9978997	.9979467	.9979928	.9980379	4.50
4.60	.9980821	.9981253	.9981677	.9982091	.9982497	.9982895	.9983284	.9983666	.9984039	.9984404	4.60
4.70	.9984761	.9985111	.9985454	.9985789	.9986117	.9986439	.9986753	.9987060	.9987361	.9987656	4.70
4.80	.9987944	.9988226	.9988502	.9988771	.9989035	.9989294	.9989546	.9989793	.9990035–	.9990271	4.80
4.90	.9990502	.9990728	.9990949	.9991166	.9991377	.9991584	.9991786	.9991983	.9992176	.9992365	4.90
5.00	.9992550–	.9992730	.9992906	.9993079	.9993247	.9993412	.9993573	.9993730	.9993884	.9994034	5.00
5.10	.9994181	.9994324	.9994464	.9994601	.9994735–	.9994865	.9994993	.9995117	.9995239	.9995358	5.10
5.20	.9995474	.9995587	.9995698	.9995806	.9995912	.9996015–	.9996116	.9996214	.9996310	.9996404	5.20
5.30	.9996495–	.9996584	.9996671	.9996756	.9996839	.9996921	.9997000	.9997077	.9997152	.9997226	5.30
5.40	.9997297	.9997367	.9997436	.9997502	.9997567	.9997631	.9997692	.9997753	.9997812	.9997869	5.40
5.50	.9997925–	.9997979	.9998033	.9998085–	.9998135	.9998185–	.9998233	.9998280	.9998326	.9998370	5.50
5.60	.9998414	.9998456	.9998497	.9998538	.9998577	.9998615	.9998653	.9998689	.9998724	.9998759	5.60
5.70	.9998792	.9998825	.9998857	.9998888	.9998919	.9998948	.9998977	.9999005–	.9999032	.9999059	5.70
5.80	.9999085–	.9999110	.9999135–	.9999159	.9999182	.9999205–	.9999227	.9999248	.9999269	.9999290	5.80
5.90	.9999309	.9999329	.9999348	.9999366	.9999384	.9999401	.9999418	.9999434	.9999450	.9999466	5.90
6.00	.9999481	.9999496	.9999510	.9999524	.9999538	.9999551	.9999564	.9999576	.9999588	.9999600	6.00
6.10	.9999612	.9999623	.9999634	.9999644	.9999655–	.9999665–	.9999675–	.9999684	.9999693	.9999702	6.10
6.20	.9999711	.9999719	.9999728	.9999736	.9999743	.9999751	.9999758	.9999765	.9999772	.9999779	6.20
6.30	.9999786	.9999792	.9999798	.9999804	.9999810	.9999816	.9999821	.9999826	.9999832	.9999837	6.30
6.40	.9999842	.9999846	.9999851	.9999856	.9999860	.9999864	.9999868	.9999872	.9999876	.9999880	6.40
6.50	.9999884	.9999887	.9999891	.9999894	.9999897	.9999900	.9999903	.9999906	.9999909	.9999912	6.50
6.60	.9999915–	.9999917	.9999920	.9999922	.9999925–	.9999927	.9999929	.9999932	.9999934	.9999936	6.60
6.70	.9999938	.9999940	.9999942	.9999944	.9999945	.9999947	.9999949	.9999950	.9999952	.9999953	6.70
6.80	.9999955–	.9999956	.9999958	.9999959	.9999960	.9999962	.9999963	.9999964	.9999965	.9999966	6.80
6.90	.9999967	.9999968	.9999970	.9999970	.9999971	.9999972	.9999973	.9999974	.9999975–	.9999976	6.90
7.00	.9999977										

TABLE 2

SUMMARY OF LEADING TABLES OF THE AREA UNDER THE NORMAL CURVE FOR ABSCISSAE EXPRESSED IN TERMS OF *P.E.*

Author	Date	Function tabulated	Number of values given	Size of interval (in P.E.)	Range of table	Number of decimals*
Encke (7)	1832	α†	357	.01 ——→ .1 ——→	.00 PE—3.40 PE 3.4 PE—5.0 PE	5
Johnson (19)	1890	α	377	.01 ——→ .1 ——→	.00 PE—3.49 PE 3.0 PE—6.1 PE	4
Rugg (30)	1917	$10{,}000(\frac{1}{2}\alpha)$	120	.05	.00 PE—5.95 PE	4 & 5‡
Pearl (26)	1930	$100(1-\alpha)$	46	.1 ——→ 1.0 ——→	1.0 PE—5.0 PE 5 PE—10 PE	Varies§
Odell (24)	1935	$\frac{1}{2}\alpha$; & $\frac{1}{2}(1-\alpha)$	73	.1‖	.0 PE—6.0 PE	4-7¶
Krause & Conrad	1937	α	701	.01	.00 PE—7.00 PE	7

*In order that the figures in this column may be directly comparable, we state the number of decimals for the function α (rather than $10{,}000(\frac{1}{2}\alpha)$ [Rugg] or $100(1-\alpha)$ [Pearl]).

†For definition of α, see text.

‡From .00 through 4.80 PE, four decimals are given; from 4.85 through 5.95 PE, five decimals. *Exceptions*: at 5.8 and at 5.9 PE, six decimals are given.

§From 1.0 through 3.8 PE, four decimals are given; from 3.9 through 4.8 PE, five decimals; at 4.9, 5, and 6 PE, six decimals; at 7 PE, seven decimals; at 8 PE, nine; at 9 PE, ten; at 10 PE, twelve.

‖Values for the following intermediate points are also given: .25 PE, .75 PE, 1.25 PE, 1.75 PE, etc., through 5.75 PE.

¶From .0 through 3.5 PE, four decimals are given; from 3.6 through 4.5 PE, five decimals; from 4.6 through 5.5 PE, six decimals; from 5.6 through 6.0 PE, seven decimals.

the abscissae $+\frac{x}{PE}$ and $-\frac{x}{PE}$ (where $\frac{x}{PE}$ *is a deviation from the mean, expressed in terms of the probable error of the distribution). The range of the table is from $\frac{x}{PE} = .00$ through $\frac{x}{PE} = 7.00$; the interval is uniformly .01; the number of decimals in the tabulated entry is uniformly seven.

A brief historical summary of leading publications of the area under the normal curve (for abscissae expressed in terms of *PE*) is given in Table 2.

CONSTRUCTION OF THE TABLE

The values of a in Table 1 were calculated by interpolation into Burgess' extensive table of $H = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$, using the relationship $1\ PE = .47693627620\ t$.† Bessel's formula, with first differences and means of second differences (**28**) was employed for values of t below 3.000 (values of x/PE below 6.30); in this portion of Burgess' table, the interval between successive values of the argument, t, is .001 (from $t = .000$ through $t = 1.500$) and .002 (from $t = 1.500$ through $t = 3.000$). At $t = 3.0$, the interval in Burgess' table jumps to .1; for this reason, third and fourth differences are of significance at and above this point, and were employed in the interpolation. All computations were carried to ten decimals. For values of H from $t = .000$ through $t = .999$, only nine decimals are given by Burgess (the values were calculated "to about twelve places, but only nine were preserved" (**2**, p. 261)); from $t = 1.000$ through $t = 3.000$, and from $t = 3.0$ through $t = 3.6$, fifteen decimals are given by Burgess, but we reduced the values to ten decimals (which is the limit of the keyboard of commercial calculating machines). The values of a in Table 1 are presented to seven decimals only, and may occasionally

*The symbol $\frac{x}{PE}$ is commonly written as $\frac{x}{r}$ in texts on the method of least squares, and in compilations of tables for physical scientists (e.g., the *Smithsonian Physical Tables* (**10**)). Other symbols have also been used — some as a matter of personal preference, some for the sake of greater brevity, and some for the sake of distinguishing between values from a small sample and values for the parent population (**4**).

†Burgess' table is found in Reference **2** (pp. 283-321). The constant, .47693627620, used to convert from *PE* to *t*, is given (to thirteen additional decimals) in Burgess' paper, on page 278.

be in error by 1 in the last place, when the seventh decimal is affected by an error in the tenth (or, possibly, the ninth) decimal.*

Whenever the eighth decimal of α was (according to our computations) 5 or greater, the seventh decimal was raised. Thus, at $x/PE = .02$, $\alpha = .01076297$ (to eight decimals), which is recorded in Table 1 as .0107630. Whenever, by this process, the seventh decimal is a raised (not an original) 5, this fact has been indicated by suffixing a minus sign to the entry in Table 1; thus, at $x/PE = .30$, the calculated value of α (to eight decimals) is .16035448; this is recorded in Table 1 as .1603545-. Analogously, at $x/PE = .32$, the value of α is recorded as .1708850- (the calculated value to eight decimals being .17088497). The suffixing of the minus sign is, of course, designed to prevent error from the progressive dropping of decimals.

In order to guard against the possibility of any constant error in the routine of computation, all calculations were performed independently by different individuals, on different calculating machines.

VERIFICATION OF THE TABLE

Within our knowledge, there are only two published sources available for the partial checking of the figures in Table 1. Pearl (**26**) gives the value of α at 7.00 *PE* as .9999977; this agrees with the value in Table 1. Odell (**24**) gives seven-decimal values of $\frac{1}{2}\alpha$ from 5.6 *PE* through 6.0 *PE*, by intervals of .1; he also gives the value of $\frac{1}{2}\alpha$ at 5.75 *PE*. All of Odell's figures agree with the (halved) values of α found by us, except at 5.8 *PE*.†

As a further check on the figures of Table 1, we have available an unpublished table of values of $\frac{1}{2}\alpha$, calculated by interpolation into

*Thus, at $x/PE = 2.96$, we found $\alpha = .9541193499$, which is recorded in Table 1 as .9541193. If the tenth figure of the computed value of α is too small by as little as .0000000001, the recorded value (to seven decimals) should be not .9541193, but .9541194. Other values of α which may be affected in the seventh place by a relatively small error in the tenth decimal are found at $x/PE = 1.29$ ($\alpha = .6157497496$), $x/PE = 2.00$ ($\alpha = .8226564499$), $x/PE = 2.85$ ($\alpha = .9454314503$), $x/PE = 3.18$ ($\alpha = .9680373498$), and at $x/PE = 3.33$ ($\alpha = .9752992501$).

†Our value for α at 5.8 *PE* is (to eight decimals) .99990848 — so that $\frac{1}{2}\alpha = .4999542$ (to seven decimals); Odell gives this value as .4999543. In an informal communication, Dr. Odell has informed us, concerning the values in his table, that "the primary source of the areas of the normal curve corresponding to various *PE* values was interpolation in Sheppard's Table. This procedure gives values which can, I believe, be relied to be in error no more than one in the seventh decimal place. To these I applied the not very objective check of graphing the entries in Sheppard's Table on a fairly large scale, smoothing the curve through them, and then determining the desired interpolated values . . . Certainly I should not want to argue very strongly for the proven correctness in the last place of the seven-decimal entries in my table, even though I believe that they are probably correct."

Sheppard's (**27**) table of $\frac{1}{2}(1 + \alpha)$.* This unpublished table (computed under the writers' direction) has the same range as Table 1, and proceeds by the same interval. From .00 *PE* through 6.67 *PE*, the values of $\frac{1}{2}\alpha$ in the unpublished table should agree reasonably well with the (halved) values of α as obtained by interpolation in Burgess' table: exact agreement is not to be looked for, since Sheppard's table itself presents only seven decimals (from .00 σ to 4.50 σ), and interpolated values cannot be expected to possess the same degree of accuracy as the original values on which they are based. As a matter of fact, from .00 *PE* through 6.67 *PE*, the unpublished figures (based on Sheppard's table) and the published figures of Table 1, agree exactly in 477 instances, and differ by only 1 in the seventh decimal, in the remaining 191 instances. From 6.68 *PE* through 7.00 *PE*, the values in the unpublished table (being based on ten-decimal values in Sheppard's table) in every case agree exactly with the figures of Table 1. This agreement between results from interpolation into Burgess' table of *H* and Sheppard's table of $\frac{1}{2}(1 + \alpha)$ is especially gratifying, since Burgess' and Sheppard's tables themselves were independently computed by their authors. Sheppard could readily have calculated his table from Burgess', but did not do so (**31**, p. 181).

ERRORS IN PREVIOUSLY PUBLISHED TABLES

Tables of the area under the normal curve, for abscissae expressed in terms of *PE*, have been rather widely published. The tables presented by Johnson (**19**), Holzinger (**16**), Pearl (**26**), and Hodg-

*The details of computation of the unpublished table of $\frac{1}{2}\alpha$ are as follows. To convert *PE*'s into units of σ, use was made of the relation, 1 PE = .6744897502 σ. To convert Sheppard's table of $\frac{1}{2}(1 + \alpha)$ into a table of $\frac{1}{2}\alpha$, .5 was subtracted from each of Sheppard's entries. Interpolation was performed by means of Bessel's formula (only the first differences and means of second differences are of significance in this interpolation). Computations were carried to eight decimals from x/PE = .00 through x/PE = 6.67, and to ten decimals from x/PE = 6.68 through x/PE = 7.00. (The shift in the number of decimals corresponds to a change in the number of decimals given in Sheppard's table.)

A word should, perhaps, be added concerning the relation 1 PE = .6744897502 σ. Burgess (**2**, p. 279) and Kondo and Elderton (**20**, p. 372) are authorities for the figure, .6744897502, which is also given by Deming and Birge (**4**, p. 125). Huntington, however, has published the relation as 1 PE = .674489749 σ (**18**, p. 11): this is too small by 1 in the last decimal. Incidentally, the 23-figure value given by Burgess (**2**, p. 279) for $\rho\sqrt{2}$ (our *PE*) appears to be in error beyond the thirteenth decimal. Multiplying Burgess' 24-figure value of ρ (viz., 47693 62762 04469 87338 3506) by the value of $\sqrt{2}$ correct to thirty decimals (viz., 1.41421 35623 73095 04880 16887 24210), we obtain 1 PE = .67448 97501 96081 74320 5179 σ — which agrees with Burgess' value for $\rho\sqrt{2}$ through the thirteenth decimal only.

man (**15**) are correct to the number of decimals given; but the majority of published tables contain one or more errors, occasionally of considerable magnitude. In the paragraphs below, we specify the errors in published tables (limiting ourselves to such tables as present four or more decimals). With minor exceptions, the order of treatment is chronological.

The table of α presented by Encke (7) and reprinted by De Morgan (5, 6) contains nine minor errors. Encke's values are too small by 1 in the last (fifth) place, at the following values of x/PE: .09, .70, 1.16, 1.27, 1.50, 2.24, 2.25, and 3.11. Encke's value at 5.00 PE is too large by 1 in the fifth place.

The table of α presented by Howe (17) contains one major error (probably a misprint), and ten minor errors. The major error occurs at .97 PE, for which α is given as .48605 — it should be .48705. Howe's minor errors are identical with Encke's (except that Howe's value at 3.11 PE is correct); in addition, Howe's values for 5.2 PE and 5.3 PE are too large by 1 in the last (fourth) place.

The first edition of the *Smithsonian Physical Tables* (13) copied Howe's major error at .97 PE, and added two other major errors (evidently misprints): the values at .04 PE and 1.37 PE were given as .02512 and .64554, respectively, when they should be .02152 and .64454, respectively. The fifth revised edition of the Smithsonian tables (8) corrected the error at .04 PE; the other two major errors were corrected in the sixth revised edition (9). All the various editions and reprints of the Smithsonian tables to date (10) contain the same minor errors as in Howe's table, through 5.0 PE; in addition the value given in the Smithsonian tables at 2.97 PE is too small by 1 in the last (fifth) place; and the values given in the Smithsonian tables for the following values of x/PE are too large by 1 in the last (fifth) place: 5.1, 5.2, 5.3, 5.4, 5.5, and 5.6.

Palmer's (25) table of the values of α contains the same errors as Howe's table (17) through 5.0 PE (including the error at .97 PE); and the same errors as in the Smithsonian tables (8), from 5.1 PE through 5.6 PE.

The sole four-figure value of $1000(\frac{1}{2}\alpha)$ given by Thorndike (33), at 5.0 PE, is too large by 1 in the last place.

The table of α presented by Benini (1) contains only one error: Benini gives the value of α at 3.70 PE as .9870 — it should be .9874.

The table of α presented by Merriman (22) contains one major error, and nineteen minor errors. The major error occurs at .97 PE — the value at this point is given as .4860; it should be .4871. Merriman's values are too small by 1 in the last (fourth) place, at the following values of x/PE: .30, .36, .58, .69, .75, 1.15, 1.24, 1.28, 1.65, 2.27, 2.29, 2.59, 2.80, 3.35, and 4.90; Merriman's values are too large by 1 in the fourth place, at the following values of x/PE: 2.33, 2.41, 5.2, and 5.3.

The table of $10{,}000(\frac{1}{2}\alpha)$ presented by Rugg (30) contains two major, and seventeen minor, errors. At .90 PE and 1.80 PE, Rugg gives the values 2291 and 3896, respectively; these should be 2281 and 3876, respectively. Rugg's values are too small by 1 in the last (fourth) place, at the following values of x/PE: .20, 1.20, 1.50, 1.95, 2.40, 2.60, 3.15, 3.20, and 4.10; they are also too small by 1 in the last (fifth) place at the following values of x/PE: 4.90, 4.95, 5.00,

5.05, and 5.35. Rugg's values are too large by 1 in the last (fourth) place at .50 PE and 1.35 PE; and too large by 1 in the last (sixth) place at 5.80 PE.

The tables of $10{,}000(\frac{1}{2}\alpha)$ presented by Garrett (11), Gregory and Renfrow (14), and Cooke (3) are identical with Rugg's table—except that these authors present four figures only (instead of Rugg's five) from 4.85 PE through 5.05 PE, and these four-figure values are all correct. The table of $\frac{1}{2}\alpha$ on page 535 of J. P. Guilford's recently published text on *Psychometric Methods* (New York: McGraw-Hill) is identical with the tables of Garrett, Gregory and Renfrow, and Cooke, except for the insertion of the decimal point.

The tables of α given by Leland (21) and by Whipple (34) contain two minor errors, the values at .30 PE and 2.8 PE in each table being too small by 1 in the last (fourth) place.

Of the six values of α given by Rietz (29), the value at 5 PE is too small by 1 in the last (sixth) place; the complement of this error occurs in Rietz' values for the function $(1 - \alpha)$.

Of the six values of α given by Gavett (12), three are incorrect. Gavett's value at 2 PE is too small by 4 in the last (fifth) place; his value at 3 PE is too small by 2 in the fifth place; his value at 5 PE is too small by 1 in the fifth place. Gavett's values of the function $(1 - \alpha)$ contain errors complementary to the ones just mentioned. Gavett also presents six values of $\frac{1}{2}\alpha$; of these, the value at 2 PE is too small by 2 in the fifth place; the values at 3 PE and 5 PE are too small by 1 in the fifth place.

The table of $\frac{1}{2}\alpha$ given by Morton (23) contains four minor errors. Morton's values at .20 PE, 1.50 PE, and 3.20 PE are too small by 1 in the last (fourth) place; the value at .50 PE is too large by 1 in the fourth place.

The table of $\frac{1}{2}\alpha$ presented by Odell (24) contains two minor errors: the value at 1.20 PE is given as .2908 — it should be .2909; the value at 5.8 PE is given as .4999543—it should be .4999542. These same errors also occur in the table of $\frac{1}{2}\alpha$ given by Broom (1a); in addition the value of $\frac{1}{2}\alpha$ at 5.9 PE is given by Broom as .4999653: it should be .4999655.

The table of 100α presented by Sorenson (32) contains one major error, and five minor errors. The major error occurs at .9 PE; the value at this point is given as 45.82 — it should be 45.62. Sorenson's values are too small by 1 in the last (fourth) place at .2 PE, .4 PE, and at .6 PE; they are too large by 1 in the fourth place at .5 PE and .8 PE. Sorenson also presents a table of $100(1 - \alpha)$; the errors in this table are complementary to those in his table of 100α.

SUMMARY

A seven-decimal table is presented of the area (α) under the unit normal curve, for abscissae expressed in terms of the "probable error" or *PE*. From the method of calculation, the partial verification by means of other tables, and the safeguards taken in the routine of computation, it is safe to consider the table reliable. Errors in previously published tables are pointed out.

REFERENCES

1. BENINI, RODOLFO. *Principii di Statistica Metodologica,* p. 230. Italy: Unione Tipografico Editrice Torinese, 1906.

1*a*. BROOM, M. E. *Educational Statistics for Beginning Students,* p. 302. New York: American Book Co., 1936.

2. BURGESS, JAMES. "On the Definite Integral $\frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$, with Extended Tables of Values." *Trans. Royal Soc. Edinburgh,* 39: 257-321 (1896-1899).

3. COOKE, D. H. *Minimum Essentials of Statistics,* p. 247. New York: Macmillan, 1936.

4. DEMING, W. E. and BIRGE, R. T. "On the Statistical Theory of Errors," p. 125. *Reviews of Modern Physics,* 6: 119-161 (1934).

5. DE MORGAN, A. "Theory of Probabilities," pp. 484-485. In *Encyclopaedia Metropolitana,* 2: 393-490 (1836 or 1837). London: Printed by William Clowes and Son, for B. Fellowes, F. and J. Rivington, etc.

6. DE MORGAN, A. *An Essay on Probabilities,* pp. xxxviii-xl. London: Printed by A. Spottiswoode, for Longman, Ormė, Brown, etc., 1838.

7. ENCKE, J. F. "Über die Methode der kleinsten Quadrate," pp. 309-312. In *Berliner Astronom. Jahrb. für 1834* (ed. by J. F. Encke). Berlin: Königl. Akademie der Wissenschaften, 1832.

8. FOWLE, F. E. (ed.). *Smithsonian Physical Tables* (fifth revised edition), p. 48. Washington, D. C.; Smithsonian Institution, 1910.

9. FOWLE, F. E. (ed.). *Smithsonian Physical Tables* (sixth revised edition), p. 57. Washington, D. C.: Smithsonian Institution, 1914.

10. FOWLE, F. E. (ed.). *Smithsonian Physical Tables* (first reprint of eighth revised edition), p. 57. Washington, D. C.: Smithsonian Institution, 1934.

11. GARRETT, H. E. *Statistics in Psychology and Education,* p. 93. New York: Longmans Green, 1926.

12. GAVETT, G. I. *First Course in Statistical Method,* p. 180. New York: McGraw-Hill, 1925.

13. GRAY, T. (ed.). *Smithsonian Physical Tables* (first edition), p. 36. Washington, D. C.: Smithsonian Institution, 1896.

14. GREGORY, C. A. and RENFROW, O. W. *Statistical Method in Education and Psychology*, p. 97. Cincinnati: C. A. Gregory Co., 1929.

15. HODGMAN, C. D. (ed.). *Handbook of Chemistry and Physics*, p. 180. Cleveland: Chemical Rubber Publishing Co., 1932.

16. HOLZINGER, K. J. *Statistical Methods for Students in Education*, p. 237. Boston: Ginn and Co., 1928.

17. HOWE, M. A. *Formulas and Tables Used in the Application of the Method of Least Squares*, pp. 24-27. Terre Haute, Ind.: Moore and Langen, 1890.

18. HUNTINGTON, E. V. "Mathematical Memoranda," p. 11. In *Handbook of Mathematical Statistics* (ed. by H. L. Rietz). Boston: Houghton Mifflin, 1924.

19. JOHNSON, W. W. *The Theory of Errors and Method of Least Squares*, p. 121. Baltimore: Press of Isaac Friedenwald, 1890.

20. KONDO, T. and ELDERTON, E. M. "Table of the Normal Curve Functions to Each Permille of Frequency." *Biometrika*, 22: 368-376 (1930-1931).

21. LELAND, O. M. *Practical Least Squares*, p. 231. New York: McGraw-Hill, 1921.

22. MERRIMAN, M. *Method of Least Squares*, p. 221. New York: John Wiley and Sons, 1909.

23. MORTON, R. L. *Statistical Tables*, p. xlvii. New York: Silver Burdett, 1928.

24. ODELL, C. W. *Statistical Method in Education*, pp. 444-445. New York: D. Appleton-Century, 1935.

25. PALMER, A. DE F. *The Theory of Measurement*, p. 226. New York: McGraw-Hill, 1912.

26. PEARL, R. *Introduction to Medical Biometry and Statistics*, p. 438. Philadelphia: W. B. Saunders, 1930.

27. PEARSON, KARL (ed.). *Tables for Statisticians and Biometricians. Part* I (second edition), pp. 2-8. London: Biometric Laboratory, University College, 1924.

28. RICE, H. L. *The Theory and Practice of Interpolation*, p. 67. Lynn, Mass.: Nichols Press, 1899.

29. RIETZ, H. L. "Random Sampling," p. 75. In *Handbook of Mathematical Statistics* (ed. by H. L. Rietz). Boston: Houghton Mifflin, 1924.

30. RUGG, H. O. *Statistical Methods Applied to Education,* p. 391. Boston: Houghton Mifflin, 1917.

31. SHEPPARD, W. F. "New Tables of the Probability Integral," pp. 182-188. *Biometrika,* 2: 174-190 (1902).

32. SORENSON, H. *Statistics for Students of Psychology and Education,* p. 366. New York: McGraw-Hill, 1936.

33. THORNDIKE, E. L. *An Introduction to the Theory of Mental and Social Measurements,* p. 149. New York: The Science Press, 1904.

34. WHIPPLE, G. C. *Vital Statistics,* p. 450. New York: John Wiley and Sons, 1923.

ANNOUNCEMENT

Research Assistantships in Mathematical Psychology

There will be available several research assistantships in mathematical psychology at The University of Chicago for the academic year 1937-1938. These assistantships are on a full-time or half-time basis. Preference will be given to graduate students in psychology who have some training in statistical theory and in under-graduate mathematics. Interested applicants should send inquiries to

L. L. THURSTONE,

The University of Chicago.

PSYCHOMETRIKA—VOL. 2, NO. 1
MARCH, 1937

ABSTRACTS OF PAPERS TO BE PRESENTED AT THE DISTRICT MEETING OF THE PSYCHOMETRIC SOCIETY, THE UNIVERSITY OF CHICAGO, ON SATURDAY, APRIL 3, 1937

Studies in the Learning Function, L. E. AND A. WILEY, Ohio Wesleyan University, Delaware, Ohio.

To the maze records of animals previously subjected to cerebral cortical injury we fitted Thurstone's theoretical learning curve

$$u = \frac{\sqrt{m}}{aK} - \frac{\sqrt{m}}{K} \cdot \frac{u}{R} . \tag{1}$$

In our analysis the curve appears in the form

$$u = A + \frac{BR}{C + R} . \tag{2}$$

A linear relationship is shown between the horizontal asymptote of the theoretical curve and the criterion which was formerly used, total-errors-minus-errors-on-the-first-trials.

Proof is given that the curve is an equilateral hyperbola which makes it possible to use the length of the semi-major axis to represent the learning situation. From this a solution for the learning constant of the animals and for the difficulty of the maze has been developed.

The properties of the vertex of the curve are discussed. At this point, all of the animals can be equated because of the fact that the first derivative is unity.

The relationship of the length of the semi-major axis to the limits of learning is discussed and a geometrical definition of insight is given. (15 minutes)

The Progress of General College Students in Mathematics, MARY L. ELVEBACK, University of Minnesota, Minneapolis, Minnesota.

An algebra test was given to a class on the first day of both the fall and winter quarters. The mean high school percentile rank was 30 and the mean C.A.T.P.R. 27: nevertheless, the mean score increased from 52 to 84 out of a possible 113 points.

Success in the course is as accurately predicted by this half-hour test as by the high school record.

Analysis of the variance reveals that students in different quarters of first score differ significantly in various measures of gain. Students of low initial status not only gain more points than those of higher initial status but their second scores are larger proportions of their first scores. In fact, the regression of

the relative gain on the first score, p, is well fit by the rectangular hyperbola

$$\frac{p_2 - p_1}{p_1} = \frac{54 - .5\,p_1}{p_1 - 2.4} .$$

Thus, as the first score increases, the relative gain decreases with a diminishing rate of decrease so that it falls off very rapidly at first but becomes almost constant for large first scores. However, the effect of initial status on the ratio of gain to possible gain is negligible except within the upper quarter of first scores. This is easily established by correlating the gain with its "most likely" value. (15 minutes)

A Rational Theory of Discrimination Learning, HAROLD GULLIKSEN AND DAEL L. WOLFLE, University of Chicago, Chicago, Illinois.

A rational theory has been developed which describes the course of discrimination learning and predicts the nature and accuracy of transfer of the discrimination to new stimuli.

The theory is sufficiently general to handle any discrimination problem, but in this paper the mathematical development will be confined to the usual type of discrimination problem in which the subject is presented with two stimuli which differ in some respects (brightness, form, size, etc.). In half of the trials these stimuli are presented spatially in an AB order, and in half of the trials in a BA order. Each order constitutes a different stimulus configuration.

We postulate the existence of two response tendencies for each stimulus configuration—to go to the right and to go to the left. The relation between the strengths of the two response tendencies for a given configuration, i, is assumed to be a function of two factors.

1. *Initial Strength*: Before any practice has been given the strength may vary due to position preferences, brightness preferences, etc.
2. *Practice*: Each reward increases the strength of a response tendency and each punishment decreases it. Practice of a response to any configuration, j, affects the strength of the same response tendency to configuration i according to a diminishing function of the distance between i and j. One such function is:

$$C_i = e^{-d^2_{ij}} \cdot k .$$

C_i = change of strength of the response tendency to configuration i, due to one practice with configuration j.
d_{ij} = distance between configurations i and j. (d_{ij} can be measured by scaling techniques.)
k = the learning constant.
If configurations i and j are identical, $d_{ij} = 0$ and $C_i = k$.

The asymptotes of the learning curve derived from these postulates depend upon the distance separating the stimulus configurations to be discriminated. The rate at which an individual curve will approach the asymptotes depends upon the subject's learning ability.

Some of the deductions regarding the ease or difficulty of learning different discrimination problems, and regarding the nature and extent of transfer to new stimuli, which have been checked against the experimental literature are:

1. If training is given under the usual conditions, learning difficulty increases as the distance between the two configurations decreases.
2. If four configurations are used in training, either an absolute or a relative discrimination can be established. The latter is easier.
3. Variation of the stimuli during training retards but does not prevent learning.
4. If subjects have been trained to make a relative discrimination, transfer to new configurations will be made on a relative basis.
5. If subjects have been trained to make an absolute discrimination, transfer to new configurations will be made on an absolute basis.
6. The accuracy of transfer will decrease as the distance between training and test configurations increases.

The derivation of the learning equation and the reasoning leading to some of the theoretical deductions will be presented. (30 minutes)

The Role of Statistical Selection of Items in Test Construction, DOROTHY C. ADKINS, The University of Chicago, Chicago, Illinois.

Statistical selection of items is distinguished from rational analysis, which may be pre-statistical or post-statistical. Those using the term "item analysis" usually mean item selection, which does not of necessity imply any rational analysis.

The application of a time-consuming selection technique, involving the interrelationships of items, to 150 items with 800 subjects is reported. Such a selection technique offered but slight improvement in predictive value over simpler methods. The low average of the 11,175 intercorrelation coefficients (.09) is consistent with this result. Data are presented to show the decrease in predictive value of a selected composite when it is tested on a new population.

That more complicated statistical techniques may be of greater value in a relatively untried field is still a possibility. Even in such a field, however, the same items cannot be repeated indefinitely. Statistical selection can be justified, therefore, only when accompanied by post-statistical analysis which provides more or better clues for future construction of items than can be attained by simpler means. That such clues do result remains to be demonstrated.

It is thus argued that pre-statistical analysis may obviate the need for post-statistical analysis and that results to date scarcely justify the application of complicated selection techniques. At the present stage, we should determine whether *any* method of item selection is of value rather than attempt to devise additional short-cuts and approximations to available statistical methods. (15 minutes)

Simple Graphic Aids for Harassed Psychometricians, HAROLD D. GRIFFIN, Nebraska State Teachers College, Wayne, Nebraska.

It is the writer's contention that we might employ certain easily constructed graphic devices more than we do, and thereby effect a saving in time and energy without reducing accuracy.

For a number of years he has used a graphic method for coding variables in multiple regression problems, generally reducing scores in all variates to a range of 20 points with 0 as low score, which is equivalent to grouping into 21

classes. This is speedier than other coding devices, and the resulting uniform range is helpful for preliminary inspection of relationships among the variables, besides facilitating rapid calculation of the various zero-order coefficients.

The writer has also found that the construction of simple graphic prediction charts aids in applying the multiple regression equation, and that these prediction charts may quite readily be adjusted by means of graphic methods to care for curvilinear tendencies in any of the original variates without employing other than the Pearson product-moment formula in obtaining zero-order correlations. (15 minutes)

The Theory of the Estimation of Test Reliability, G. F. KUDER AND M. W. RICHARDSON, University of Chicago, Chicago, Illinois.

Beginning with a precise definition of the equivalence of test forms, the reliability coefficient is expressed in its theoretically best form. A series of special assumptions are made for cases in which item analyses of varying degrees of completeness are available. The resulting formulas represent various degrees of approximation to the theoretical coefficient. The last of these,

$$r_{tt} = \frac{n^2\sigma^2{}_t + M^2 - Mn}{n(n-1)\sigma^2{}_t},$$

gives a rapid method of estimating the reliability coefficient from the commonly computed parameters of the test score distribution.

An application of several of the formulations is made to test data, in order to illustrate the degree of approximation involved in each method. (30 minutes)

The Estimation of Factors: A-Five Methods Applied to a Bi-Factor Problem, FRANCES SWINEFORD, University of Chicago, Chicago, Illinois.

The ultimate goal of the factor analyst is the measurement of factors for individuals. In case the factors have been described in terms of a large number of tests, the labor of calculating regression estimates and applying them to individual scores is so great as to be out of the question in practice.

It is our purpose to outline a number of short-cuts, and to apply them to actual data. A small set of twelve tests is employed in order that the complete regression estimate may be used as a standard with which to compare the short-cut estimates.

The comparisons are made on the basis of the multiple correlation coefficients, and of the intercorrelations among the estimates. (15 minutes)

The Estimation of Factors: B—Systems of Regression Equations for the Estimation of Factors, HARRY H. HARMAN, University of Chicago, Chicago, Illinois.

In this paper we set up five methods of estimating an individual's factor abilities from a given factor solution. The regression equations for the estimation of each factor are of several types: those involving every test, those involving only some of the tests, and those involving groups of tests considered as unit tests.

These methods are used in the estimation of factors obtained by four different factor analyses of an hypothetical problem. We compare the multiple correlation coefficients of the different factor estimates for a given set of factors to

obtain the most practical and statistically sound method. Then we compare the accepted estimates of factors obtained from different factorial solutions to see how nearly the estimates of factors follow the same laws of comparison as the actual factors. (15 minutes)

A Method of Factor Analysis by Means of which all Co-ordinates of the Factor Matrix are given simultaneously, PAUL HORST, The Proctor and Gamble Company, Cincinnati, Ohio.

In general, the methods of factor analysis developed during the past five years are based on the reduction of the correlational matrix by successive steps. The first factor loadings are determined and eliminated from the correlational matrix, giving a residual matrix. This process is continued for successive factor loadings until the elements of the last obtained residual matrix may be regarded as due to chance.

The method outlined in this paper does not involve the progressive disintegration of the correlational matrix. Letting r be the postulated number of factors and n the order of the correlational matrix, we calculate $r + 1$ vectors, where the $(i+1)$'th vector is equal to the i'th vector premultiplied by the correlational matrix. From these vectors certain vector equations are derived which involve only the roots of the correlational matrix. The factor matrix is then expressed as a function of the $(r+1)$ vectors and the r roots. The solution for the roots is not linear, but once these are obtained the solution for the factor matrix is linear.

It is possible to solve for the basic vectors directly from the score matrix without first calculating the correlational matrix. The method is well adapted to the use of Hollerith equipment. If the variables are dichotomous, the calculation of the basic vectors reduces to a simple sorting and tabulating routine, and several hundred variables or more may be easily manipulated.

The mathematical theory underlying the method is developed, and illustrative material is presented. (30 minutes)

Matrix Approximation Criteria, GALE YOUNG, University of Chicago, Chicago, Illinois.

Linear factor analysis requires approximating given data matrices in such way as to materially reduce the rank without introducing too serious error. Some criteria for estimating the seriousness of this error are considered:

(*A*) *Size of error*: Let $\alpha(A,B)$ be a scalar measure of the error in replacing A by B; and, for a given A, let $\phi(r)$ be the minimum value of α for B of rank r. ϕ is then the minimum error possible in reducing the rank to r; and consideration of ϕ against r shows if there is a value $\overline{r}$ for r significantly indicated. This is an empirical matter; the matrix A may or may not be susceptible to such rank reduction.

Some forms for α are listed. For some of these the minimizing problem is solved, for some it is not. One of the forms is given in connection with a precise definition for "communalities".

(*B*) *Form of error*: Besides being small, it is desirable that the residuals exhibit some degree of randomness. A characterization of an ideally random

matrix is suggested, although the probability significance of departure therefrom is not yet worked out. (15 minutes)

Current Misuse of the Factorial Methods, L. L. THURSTONE, University of Chicago, Chicago, Illinois.

Those psychologists who have devoted themselves recently to the development of factor analysis should be encouraged by the very general interest in these new methods. The number of papers in the psychological journals that involve factor analysis is increasing rapidly. It is therefore a serious matter that the large majority of these papers involve misinterpretation of the factorial methods. Factor theory is still imperfect and there are some very challenging theoretical problems for the mathematicians to solve in making these analytical tools even more powerful than they now are. But if the misapplications of factor methods continue at the present rate, we shall soon find general disappointment with the results because they are usually meaningless as fas as psychological interpretation is concerned.

With the hope of assisting in the best use of these new methods in their present state of development we may summarize the most common reasons for the meaningless results in most of the current factorial studies of psychological problems.

1) The number of psychologically basic factors in a battery of tests must be considerably smaller than the number of tests in the battery.

2) The diagonals must be regarded as unknown communalities unless one can be certain that specific factors are absent.

3) No matter how the correlational matrix is factored, the axes must be rotated into a simple configuration before any psychological interpretation can be made. The frequent attempts to find psychological interpretation for the centroid axes without rotation and for the principal components without rotation are examples of this frequent error.

4) No meaningful component can be identified unless each factor is overdetermined with three or four or more tests. Most of the current studies are made on batteries that are far too short.

5) In order that meaningful factors shall emerge it is essential that the individual tests be as simple as possible as regards factorial composition. If all of the tests involve many psychological factors, then the identification of the factors becomes very difficult. Composite tests like the Stanford Binet and composite group tests of intelligence are so complex in factorial composition that they are not useful in the identification of basic factors unless they are analyzed by separate items.

While the factorial methods are not yet perfected, they are sufficiently developed so that they can be used in the solution of many fundamentally important psychological problems. The factorial methods must be so formulated that they satisfy not only the requirements of mathematical rigor but also the restrictions of the psychological problems which they are intended to solve. Some of the conditions here listed are not generally accepted but the writer is quite certain that psychologically meaningful results will not be obtained unless we satisfy these conditions. The practical answer to any controversy about the conditions here listed will be in terms of psychological results. These conditions refer to the present state of knowledge of factor theory. (30 minutes)